Editorial

Ecography 33: 985–989, 2010
doi: 10.1111/j.1600-0587.2010.06953.x

New trends in species distribution modelling

Niklaus E. Zimmermann, Thomas C. Edwards Jr., Catherine H. Graham, Peter B. Pearman and Jens-Christian Svenning

N. E. Zimmermann (niklaus.zimmermann@wsl.ch) and P. B. Pearman, Land Use Dynamics, Swiss Federal Research Inst. WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland. – T. C. Edwards, U.S. Geological Survey, Utah Cooperative Fish and Wildlife Research Unit and Dept of Wildland Resources, Utah State Univ., Logan, UT 84322-5290, USA. – C. H. Graham, Dept of Ecology and Evolution, Stony Brook Univ., NY 11794, USA. – J.-C. Svenning, Ecoinformatics and Biodiversity Group, Dept of Biological Sciences, Aarhus Univ., Ny Munkegade 114, DK-8000 Aarhus C., Denmark.

Species distribution modelling has its origin in the late 1970s when computing capacity was limited. Early work in the field concentrated mostly on the development of methods to model effectively the shape of a species' response to environmental gradients (Austin 1987, Austin et al. 1990). The methodology and its framework were summarized in reviews 10–15 yr ago (Franklin 1995, Guisan and Zimmermann 2000), and these syntheses are still widely used as reference landmarks in the current distribution modelling literature. However, enormous advancements have occurred over the last decade, with hundreds – if not thousands – of publications on species distribution model (SDM) methodologies and their application to a broad set of conservation, ecological and evolutionary questions. With this special issue, originating from the third of a set of specialized SDM workshops (2008 Riederalp) entitled 'The Utility of Species Distribution Models as Tools for Conservation Ecology', we reflect on current trends and the progress achieved over the last decade.

SDMs represent an empirical method to draw statistical inferences about the drivers of species' ranges under different conservation, ecological and evolutionary processes. Many efforts have been made to improve the statistical bases of SDMs, including: 1) the implementation of new statistical methods (Yee and Mitchell 1991, Phillips et al. 2006, Prasad et al. 2006, Elith et al. 2008); 2) the evaluation of sampling design on model performance (Edwards et al. 2006, Guisan et al. 2006); 3) the exploration of sample size and prevalence on the accuracy of SDMs (Stockwell and Peterson 2002, Zimmermann et al. 2007, Wisz et al. 2008); 4) the removal of spatial autocorrelation from model fitting (Dormann et al. 2007); 5) the comparison of a range statistical methods for SDMs (Elith et al. 2006), or 6) the evaluation of models (Allouche et al. 2006, Smulders et al. 2010), to list but a few. In summary, a large number of published papers have contributed to this methodological advancement.

Over the last decade, many publications have additionally highlighted unresolved issues in SDM (Austin 2002, 2007, Pearson and Dawson 2003, Guisan and Thuiller 2005, Araújo and Guisan 2006, Elith and Graham 2009, Elith and Leathwick 2009). Many of these issues, such as the clarification of the niche concept, model parameterization schemes, model selection, model evaluation and variable selection methods have since been explored, albeit with mixed success, or are under active research today. These constitute the current trends in SDMs, and represent, to some degree, a movement away from methodological studies towards the incorporation and use of SDMs in addressing conservation, ecological, and evolutionary questions.

A more limited number of studies have employed SDMs for resolving crucial conservation issues, or exploring macroecological and evolutionary concepts and theory. Examples of the latter are the analysis of limited range filling by Svenning and Skov (2004), of niche properties and range size on sensitivity to climate change (Thuiller et al. 2005a), of the effect of climatic vs biotic factors on species ranges (Araújo and Luoto 2007), or of the effect of climatic extremes on species ranges (Zimmermann et al. 2009). In addition, several important contributions have been made towards clarifying theoretical aspects around the relationship of SDMs and the niche concept (Austin 2002, 2007, Guisan and Thuiller 2005, Soberon 2007).

Here, we provide an overview of nine papers originating from the 2008 Riederalp workshop that contribute significantly to these developing trends in use and application of SDMs. The papers focus on five active areas of research involving SDMs, including: 1) historical legacies; 2) niche stability and evolution; 3) biotic interactions; 4) the importance of sample designs; and 5) species invasions. We believe these papers set the stage for future SDM research questions, and represent several next logical steps in SDM research and application.

Current trends – the workshop papers

The legacy of history

One of the fundamental assumptions in SDM is that the current range of species is in equilibrium with the explanatory variables that drive their ranges. As a result, the effect of history on range size and distribution patterns is generally not considered. This assumption of equilibrium has been criticized repeatedly in the past, and several authors have demonstrated that conclusions made under this assumption can be wrong (Pearson and Dawson 2003). In this issue, two papers explore how historical factors influence current species richness patterns (Svenning et al. 2010), and how historical climate variability and associated hypothesized range dynamics influence observed patterns (Graham et al. 2010). Both papers evaluate current patterns in the light of the legacy of history. Svenning and co-workers use SDM and species richness modeling to estimate realized (R, actual) to potential (P, according to site conditions) richness ratios in tree assemblages across Europe, and to assess the potential drivers of these ratios. They find support for a negative effect of geographic fragmentation, a positive effect of accessibility to postglacial recolonization, and a positive effect of topographic heterogeneity on R/P ratios. The fragmentation and accessibility effects suggest that dispersal constitutes a strong influence on R/P ratios. Graham and co-workers explore how historical ranges, considered as shifting refugia through time, serve to explain the current patterns of species richness in the Australian Wet Tropics. While others have considered how past climate influences current patterns of diversity, Graham and co-workers developed a new approach that quantifies how connected a given range has been. Essentially, the approach resembles a least-cost analysis to quantify connectivity through time. Analyses that combine static and dynamic elements likely will see increased use in the future of SDM analyses and applications. The increasing availability of historical climate information should lead to increased understanding about the effect of history on current ranges and on species, phylogenetic and functional diversity.

Niche stability and evolution

The analysis of niche evolution is an active area of ecological and evolutionary research (Wake et al. 2009). One way to evaluate niche stability is to analyze environmental constraints over a phylogenetic tree and to explore to what degree specified environmental associations can be interpreted as local adaptation. This idea may be evaluated by applying SDM to ecologically defined populations rather than to species, i.e. by examining the niche response to environmental drivers at a sub-species level. Using this approach, Pearman and co-workers (2010) find that sub-species ranges collectively do not coincide with the species range, indicating that niches fitted with data from the full taxon can over- or underpredict distributions under certain conditions. More importantly, they find that the shapes of environmental response curves generated in sub-species models differ from those that emerge at the species level. This paper is an important step towards a better understanding of niche shape and the effects of niche evolution and adaptation to environmental constraints.

Biotic interactions

Most SDMs are calibrated under the assumption that biotic interactions do not influence species range patterns (Huntley et al. 1995, Bakkenes et al. 2002), or only affect patterns at small spatial scales (Pearson and Dawson 2003, Dormann et al. 2007, Heikkinen et al. 2007). As a consequence, models are often built using abiotic predictors alone. However, examples exist to demonstrate how the incorporation of biotic interactions into SDMs better models species distributions and responses to environmental change, such as disturbance or climate (Araújo and Luoto 2007, Meier et al. 2010a). Biotic interactions have been included in SDMs as occurrence (Heikkinen et al. 2007), counts or frequencies (Leathwick and Austin 2001, Leathwick 2002, Heikkinen et al. 2007), proportional data (Meier et al. 2010a), basal area (Rouget et al. 2001), or as a competition coefficient (Strubbe et al. 2010). From a theoretical viewpoint, the response of species along macroclimatic gradients generally represents the realized ecological niche sensu Hutchinson (Austin et al. 1990, Araújo and Guisan, 2006). As a consequence, large-scale species patterns are clearly influenced by both abiotic predictors and biotic interaction variables. The importance of biotic interactions may vary according to scale and position along environmental gradients (Meier et al. 2010a). Here, Meier and co-workers (2010b) evaluate the degree to which adding biotic predictors helps to explain tree species distributions, and whether biotic and abiotic predictors explain similar or dissimilar proportions of the variability in species patterns. Interestingly, they find surprisingly little overlap in the proportion of explained model variance between the two predictor sets, and they found a linkage between the importance of biotic predictors and functional traits. In a similar study, Pellissier and co-workers (2010) explore the effect of one dominant dwarf shrub species as predictor of the distribution of other arctic-alpine species. Again, surprising little overlap exists in the predictive power of the abiotic and biotic predictor sets. Further, plant functional traits of modeled species partially explain the degree to which addition of biotic predictors enhances SDM predictive power.

The importance of design for sampling

Design and analysis are tightly linked elements of the statistical models used in SDMs. Yet, SDMs are often built from museum records or from large and inconsistently sampled databases having, at best, limited underlying designs. Differences in design quality can strongly influence the reliability of SDMs (Edwards et al. 2006). However, the high complexity of large study areas renders many conceptually derived designs unsolvable, or at least very costly to achieve. As a result, many data sets used in SDMs contain biases ranging from incomplete sampling of environmental and geographic space to targeted rather than random selection of sampling units. It is therefore

important to explore different approaches to design and to study their effect on model performance and bias. Such knowledge allows us to avoid pitfalls in modeling and in using models to inform the questions under study. Albert and co-workers (2010) implemented a set of sampling strategies in a semi-virtual experiment with known reality (designed response) to demonstrate how designs of different complexity yield models of varying performance and bias. Notably, a distinction is made between different sampling spaces, a concept that is important when optimizing a sampling in complex landscapes. Le Lay and co-workers (2010) demonstrate the capacity of model-based sampling to detect populations and to develop SDMs of rare species in Switzerland. In addition to using two different statistical methods, the models are developed at two different spatial grain sizes (50 m and 1 km). Combined models using both grain size and methods yield greater species detectability and model performance. The latter indicates that drivers and response often operate at more than one scale.

Species invasion

SDMs have often been used to assess the invasion potential of species into new habitats (Beerling et al. 1995, Peterson 2003, Thuiller et al. 2005b) or to explore niche stability and niche modeling issues during invasion (Broennimann et al. 2007, Fitzpatrick et al. 2007, Beaumont et al. 2009). Several key problems arise from the analysis of the invasive potential of species to new continents. On one hand, sampling the currently invaded range may underestimate the niche response, because the species may not yet realize its full niche potential. On the other hand, fitting a species from data of the original range may not yield the full niche either, as the original range may not exhibit all possible climate and site combinations that could possibly be inhabited by a species. In addition, invaded ranges usually go along with changes in biotic interactions, and therefore we expect changes in the realized niche during invasions. Comparing original and invaded ranges therefore represents an interesting challenge, although it is prone to misunderstandings unless studied in a careful, systematic fashion. Despite these challenges, study of invasive species is an important aspect of model-based management, as potential damages from invasives are high and very costly. Richardson and co-workers (2010) provide a study that optimizes the management of an invasive species in South Africa by combining SDM with a dynamic cellular automata model under projected future climates. Using this approach, the authors are able to prioritize areas of major concern for potential future spread of the invasive species. Vicente and co-workers (2010) explore a range of different drivers of invasion by analyzing a large set of invasive plants in Portugal. Instead of simply analyzing many potential predictors, they apply a nested model approach, starting with a climate model and then nesting models with additional landscape, land use (incl. intensity) and structural predictors. By this, they are able to demonstrate that climate exerts a primary role, while other landscape variables are additionally important where climate is favorable for predicting invasive plant species patterns and resulting alien biodiversity. They find, however, differences among CSR strategy types in their response to the nested drivers.

Conclusions

The five themes identified in this special issue will likely remain important elements of SDM research over the coming years. We specifically expect novel solutions from combining SDM with other, more dynamic or more physiological approaches, such as that presented by Richardson et al. (2010). Such approaches have the potential to provide workable solutions to important conservation, ecological and evolutionary questions. Without being exhaustive, we believe a) that aspects of niche evolution will become a central focus as SDMs and phylogenetic research mutually benefit from each other; and b) that advances in dynamic modeling linked with SDMs and conceptual or theoretical advances related to community assembly rules will likely exert a strong influence on SDMs and species modeling.

Niche evolution, phylogeographic and phylogenetic research together with increased availability of historical climate maps and the building of large databases of species distributions will allow novel analysis of both historical and contemporary drivers of species ranges, and will likely provide us with novel capacities to improve our projections of species ranges into the future. We therefore expect significant advances from these fields for our understanding of the general drivers of species ranges. On the other hand, SDMs still performs rather weakly at small spatial scales. Improvements in this domain require the development of applicable community assembly and species filtering rules and approaches. Such rules likely build on both phylogenetic and functional elements, and new approaches can be expected to further our understanding how functional traits affect community structures and biotic interactions.

Acknowledgements – We acknowledge financial support by the Swiss Federal Research Institute WSL for organizing the Third Riederalp Workshop in 2008. NEZ, PBP and TCE were funded by the 6th and 7th European Framework Programme Grant GOCE-CT- 2007-036866 (ECOCHANGE) and ENV-CT-2009-226544 (MOTIVE). JCS was funded by the Danish Council for Independent Research–Natural Sciences (grant 272-07-0242). Mention of a product by name does not constitute endorsement by the U.S. Geological Survey.

References

Albert, C. H. et al. 2010. Sampling in ecology and evolution bridging the gap between theory and practice. – Ecography 33: 1028–1037.

Allouche, O. et al. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). – J. Appl. Ecol. 43: 1223–1232.

Araújo, M. B. and Guisan, A. 2006. Five (or so) challenges for species distribution modelling. – J. Biogeogr. 33: 1677–1688.

Araújo, M. B. and Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. – Global Ecol Biogeogr. 16: 743–753.

Austin, M. P. 1987. Models for the analysis of species' response to environmental gradients. – Vegetatio 69: 35–45.
Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. – Ecol. Model. 157: 101–118.
Austin, M. P. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. – Ecol. Model. 200: 1–19.
Austin, M. P. et al. 1990. Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. – Ecol. Monogr. 60: 161–177.
Bakkenes, M. et al. 2002. Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. – Global Change Biol. 8: 390–407.
Beaumont, L. J. et al. 2009. Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. – Divers. Distrib. 15: 409–420.
Beerling, D. J. et al. 1995. Climate and the distribution of *Fallopia japonica* – use of an introduced species to test the predictive capacity of reponse surfaces. – J. Veg. Sci. 6: 269–282.
Broennimann, O. et al. 2007. Evidence of climatic niche shift during biological invasion. – Ecol. Lett. 10: 701–709.
Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – Ecography 30: 609–628.
Edwards, T. C. et al. 2006. Effects of sample survey design on the accuracy of classification tree models in species distribution models. – Ecol. Model. 199: 132–141.
Elith, J. and Graham, C. H. 2009. Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. – Ecography 32: 66–77.
Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. – Annu. Rev. Ecol. Evol. Syst. 40: 677–697.
Elith, J. et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. – Ecography 29: 129–151.
Elith, J. et al. 2008. A working guide to boosted regression trees. – J. Anim. Ecol. 77: 802–813.
Fitzpatrick, M. C. et al. 2007. The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? – Global Ecol. Biogeogr. 16: 24–33.
Franklin, J. 1995. Predictive vegetation mapping: geographical modelling of biospatial patterns in relation to environmental gradients. – Prog. Phys. Geogr. 19: 474–499.
Graham, C. H. et al. 2010. Dynamic refugia and species persistence: tracking spatial shifts in habitat through time. – Ecography 33: 1062–1069.
Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – Ecol. Model. 135: 147–186.
Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.
Guisan, A. et al. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. – Ecol. Model. 157: 89–100.
Guisan, A. et al. 2006. Using niche-based models to improve the sampling of rare species. – Conserv. Biol. 20: 501–511.
Heikkinen, R. K. et al. 2007. Biotic interactions improve prediction of boreal bird distributions at macro-scales. – Global Ecol. Biogeogr. 16: 754–763.
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.
Le Lay, G. et al. 2010. Prospective sampling based on model ensembles improves the detection of rare species. – Ecography 33: 1015–1027.
Leathwick, J. R. 2002. Intra-generic competition among Nothofagus in New Zealand's primary indigenous forests. – Biodivers. Conserv. 11: 2177–2187.
Leathwick, J. R. and Austin, M. P. 2001. Competitive interactions between tree species in New Zealand's old-growth indigenous forests. – Ecology 82: 2560–2573.
Meier, E. S. et al. 2010a. Co-occurrence patterns of trees along macro-climatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica* L. – J. Biogeogr. doi:10.1111/j.1365-2699.2010.02405.x
Meier, E. S. et al. 2010b. Biotic and abiotic variables show little redundancy in explaining tree species distributions. – Ecography 33: 1038–1048.
Pearman, P. B. et al. 2010. Within-taxon niche structure: niche conservatism, divergence and predicted effects of climate change. – Ecography 33: 990–1003.
Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – Glob. Ecol. Biogeogr. 12: 361–371.
Pellissier, L. et al. 2010. Species distribution models reveal apparent competitive and facilitative effects of a dominant species on the distribution of tundra plants. – Ecography 33: 1004–1014.
Peterson, A. T. 2003. Predicting the geography of species' invasions via ecological niche modeling. – Q. Rev. Biol. 78: 419–433.
Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – Ecol. Model. 190: 231–259.
Prasad, A. M. et al. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. – Ecosystems 9: 181–199.
Richardson, D. M. et al. 2010. Accommodating scenarios of climate change and management in modelling the distribution of the invasive tree *Schinus molle* in South Africa. – Ecography 33: 1049–1061.
Rouget, M. et al. 2001. Determinants of distribution of six *Pinus* species in Catalonia, Spain. – J. Veg. Sci. 12: 491–502.
Smulders, M. et al. 2010. A spatially explicit method for evaluating accuracy of species distribution models. – Divers. Distrib. 16: 996–1008.
Soberon, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. – Ecol. Lett. 10: 1115–1123.
Stockwell, D. R. B. and Peterson, A. T. 2002. Effects of sample size on accuracy of species distribution models. – Ecol. Model. 148: 1–13.
Strubbe, D. et al. 2010. Assessing the potential impact of invasive ring-necked parakeets *Psittacula krameri* on native nuthatches *Sitta europeae* in Belgium. – J. Appl. Ecol. 47: 547–557.
Svenning, J.-C. and Skov, F. 2004. Limited filling of the potential range in European tree species. – Ecol. Lett. 7: 565–573.
Svenning, J.-C. et al. 2010. Geography, topography, and history affect realized-to-potential tree species richness patterns in Europe. – Ecography 33: 1070–1080.
Thuiller, W. et al. 2005a. Niche properties and geographical extent as predictors of species sensitivity to climate change. – Global Ecol. Biogeogr. 14: 347–357.
Thuiller, W. et al. 2005b. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. – Global Change Biol. 11: 2234–2250.
Vicente, J. et al. 2010. What drives invasibility? A multi-model inference test and spatial modelling of alien plant species

richness patterns in northern Portugal. – Ecography 33: 1081–1092.
Wake, D. B. et al. 2009. Biogeography, changing climates, and niche evolution. – Proc. Natl Acad. Sci. USA 106: 19631–19636.
Wisz, M. S. et al. 2008. Effects of sample size on the performance of species distribution models. – Divers. Distrib. 14: 763–773.
Yee, T. W. and Mitchell, N. D. 1991. Generalized additive-models in plant ecology. – J. Veg. Sci. 2: 587–602.
Zimmermann, N. E. et al. 2007. Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. – J. Appl. Ecol. 44: 1057–1067.
Zimmermann, N. E. et al. 2009. Climatic extremes improve predictions of spatial patterns of tree species. – Proc. Natl Acad. Sci. USA 106: 19723–19728.