

Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions

Damaris Zurell^{1,2}*, Jane Elith³ and Boris Schröder^{1,4}

¹Institute of Earth and Environmental Sciences, University of Potsdam, Karl-Liebknecht-Str. 24/25, D-14476 Potsdam, Germany, ²Institute for Biochemistry and Biology, University of Potsdam, Maulbeerallee 2, D-14469 Potsdam, Germany, ³School of Botany, University of Melbourne, Parkville, Vic. 3010, Australia, ⁴Landscape Ecology, Technische Universität München, Emil-Ramann-Straße 6, D-85354 Freising, Germany

*Correspondence: Damaris Zurell, Institute for Biochemistry and Biology, University of Potsdam, Maulbeerallee 2, D-14469 Potsdam, Germany.

E-mail: damaris.zurell@uni-potsdam.de

ABSTRACT

Data limitations can lead to unrealistic fits of predictive species distribution models (SDMs) and spurious extrapolation to novel environments. Here, we want to draw attention to novel combinations of environmental predictors that are within the sampled range of individual predictors but are nevertheless outside the sample space. These tend to be overlooked when visualizing model behaviour. They may be a cause of differing model transferability and environmental change predictions between methods, a problem described in some studies but generally not well understood. We here use a simple simulated data example to illustrate the problem and provide new and complementary visualization techniques to explore model behaviour and predictions to novel environments. We then apply these in a more complex real-world example. Our results underscore the necessity of scrutinizing model fits, ecological theory and environmental novelty.

Keywords

Environmental niche, extrapolation, inflated response curves, novel environment, sampling space, species distribution models.

INTRODUCTION

Predictive species distribution models (SDMs, Guisan & Zimmermann, 2000; Elith & Leathwick, 2009) have become a prominent technique in conservation biogeography and are increasingly used as prediction tools for environmental change forecasts and invasive species research (Franklin, 2010). Numerous SDM algorithms exist with varying degrees of model complexity (Elith et al., 2006; Heikkinen et al., 2006). Several studies have shown that these algorithms can predict substantially different future potential ranges even if current predictions are largely congruent (Thuiller, 2004; Buisson et al., 2010). Explanations for varying behaviour usually point to the extent to which the environmental range was covered by the training data and to the specific assumptions made by each algorithm when extrapolating beyond that range (Thuiller et al., 2004; Pearson et al., 2006; Elith & Graham, 2009). Williams & Jackson (2007) argued that data limitations may impede extrapolation to novel environments because the species' niche may not be fully represented by data (here, termed 'truncated niches') and, depending on the direction of environmental change, currently unobserved portions of the niche may open up. Fitzpatrick & Hargrove (2009) contended that predictions should not be attempted to environmental conditions without analogues to the combinations under which the model was calibrated, or at least that maps should indicate where extrapolation has occurred.

Useful ideas are emerging for probing models and predictions, enabling users to understand model behaviour in novel space. For instance, environmental spaces have been compared using principal component analyses and metrics summarizing differences between niches (Broennimann et al., 2007; Warren et al., 2008; Medley, 2010); impacts of sample design on environmental and niche coverage have been explored and related to models and their predictions (Albert et al., 2010); and methods for mapping novel environments in geographic space have been suggested (Williams et al., 2007; Platts et al., 2008; Elith et al., 2010). Here, we add to these by focussing on the issue of combinations of variables that are within the sampled range of each predictor treated individually, but are nevertheless outside of the sampled environmental space (Fig. 1, hatched areas). These tend to be overlooked in visualization methods (cf. Fitzpatrick & Hargrove, 2009). For instance, partial dependence functions (i.e. plots of the fitted functions that show the effect of a variable on the response after accounting for the average effects of all other variables in the model) are plotted along the full gradient of each variable represented in the data, regardless of the coverage along that



Figure 1 Conceptual diagram illustrating three situations how species niche may be represented in sampled environmental space (dotted ellipse): (1) a species niche is entirely represented by sample space (species 1) (2) the niche is 'truncated' because samples do not exist for part of one or more environmental gradients (species 2) and (3) the edge of the niche abuts the edge of the sampled space, and no samples exist beyond it (species 3). The hatched square represents the 'implied' sample space that is implicitly assumed to be known when focussing on the sampled, univariate ranges of all environmental predictors individually instead of explicitly focussing on the multivariate combinations of environmental predictors represented in the sample.

gradient of other environmental dimensions. MaxEnt's multivariate environmental similarity surface (MESS, Elith *et al.*, 2010) takes a related box-like or envelope viewpoint by analysing environmental coverage one variable at a time and reporting as novel those conditions outside the environmental hyper-dimensional rectangle. However, not all multivariate combinations of the environmental conditions may be represented in the data. We define those parts of the environmental space that are within that box but nevertheless outside the sample space as 'implied sample space' (hatched areas of Fig. 1). Here, we show that existing methods can fail to clarify why predictions differ, and we provide new and complementary visualization techniques that will be relevant for many species modelling problems.

DEMONSTRATING PREDICTION PROBLEMS: SIMULATED SPECIES

Figure 1 illustrates three situations that can arise when sampling in geographic space (Williams & Jackson, 2007; Albert *et al.*, 2010). For species 2 and 3, no samples exist for parts of the environmental niche or for the niche edges. These may not be problematic if the intention is simply to model the distribution of that species in the sampled space, but as soon as models to these data are used for prediction to new times and places which might contain environments outside of the training sample, difficulties arise.

To simulate data representing the situations of Fig. 1, a virtual species (Zurell *et al.*, 2010) was created (using logistic

regression) that exhibited a unimodal response to temperature and a positive linear response to percent woodland cover (Fig. 2a; for details see Appendix S2 in Supporting Information). The entire simulation study was built in R (R Development Core Team, 2010), and we provide code in Appendix S1. For each situation, 1000 samples were drawn and converted to binary observations by using the simulated response (varying from 0 to 1) as the success rate for one sample of the binomial distribution. For species 1, samples cover the entire environmental space, while for species 2 (truncated niche), the samples cover the full univariate range of each environmental variable individually, but combinations of the two are missing (Fig. 2a). SDMs were fitted to these samples using generalized additive models (GAMs) with cubic smoothing splines, four degrees of freedom and no interactions, and boosted regression trees (BRTs) with tree complexity of 1 (tree stumps; note that in our examples higher tree complexity results in similar extrapolation behaviour). We chose these methods as examples of the range of current methods, spanning standard regression techniques to advanced machine learning methods (for overviews see Elith et al., 2006; Heikkinen et al., 2006). The models were then used to predict across the full environmental space spanned by the environmental gradients of the individual predictors, meaning that for species 2, predictions were made to new combinations of variables.

For species 1 (entire niche sampled), both methods were successful in fitting the true response (Fig. S1). Because the environmental niche of the species was truncated in the training data for species 2, predictions for the unsampled combinations required extrapolation. As a result of the way our cubic splines and regression trees extrapolate, GAM continued the fitted trend to 'unknown' sites, while BRT predicted a constant value from the last 'known' site leading to inaccurate model predictions in those parts of the unsampled environment space with high woodland cover, and particularly those that also have lower or higher than optimal temperatures (Fig. 2d; Fig. S2). The latter is not obvious from the usual partial dependence plots (Fig. 2b) because these are derived at average values of other predictors, for which this model performs reasonably well. Similar extrapolation errors also occur if niche edges coincide with the limits of the recorded environmental space (species 3; Fig. S3).

NEW TOOLS FOR VISUALIZATION

The simulation study was simple, and use of three-dimensional plots (e.g. Fig. 2d) was sufficient to demonstrate the model fit and its implications for predictions to unsampled combinations of predictors (cf. Fig. S2). In most situations, though, models have more than two covariates, and predictions are also mapped. Hence, we suggest two new tools that will highlight predictions to new combinations of variables.

First, we propose to 'inflate' conventional response curves (partial dependence plots) by visualizing the effects of all variables in the model over their full range, and at the same time plotting the available data in that space. Basically, inflated



Figure 2 Simulated data example for species 2 with truncated niche. (a) True response surface. (b) Partial dependence plots for generalized additive model (GAM) and boosted regression tree (BRT). (c, d) show fitted response surfaces for GAM and BRT. Grey dots at P = 0 in panels (a, c, d) represent sampled absences, and black dots at P = 1 represent sampled presences.

response curves are an abstracted 2D version of multidimensional response surfaces. These show the effect of a variable on the response while accounting not only for the average effects of the other variables but also for minimum and maximum (and median and quartile) values. Thus, the response plot for any one variable consists of many response curves representing all possible combinations of all other variables in the model (for code see Appendix S1; for detailed description see Appendix S3). Because the number of combinations grows exponentially with the number of variables and restricts computational feasibility, we use Latin hypercube sampling to reduce dimensionality for large numbers of variables. This is simply a means to efficiently sample a representative subset from all possible combinations of environmental predictors (Carnell, 2009).

Second, we propose to extend the idea of MESS maps by not only focussing on the environmental range of predictors individually but also on combinations of environmental predictors. By that, we are able to identify those parts of the environmental space that are within the sampled, univariate range of the individual predictors but nevertheless represent new multivariate combinations of these ('implied sample space' of Fig. 1). This 'environmental overlap' (or 'environmental gap' if one wants to emphasize that certain parts of the prediction space may not be represented in the sample space) can be determined by splitting the training or reference data into a specified number of bins where each bin holds a unique combination of environmental predictor values. Any bins in test or prediction data that do not overlap with these reference bins are defined as novel environments. An environmental overlap mask can be used to highlight predictions where the model must extrapolate to novel environments (cf. 'null prediction' in Fitzpatrick & Hargrove, 2009), for example, within inflated response curves and in prediction maps (for code see Appendix S1; for detailed method description see Appendix S3). Note that a bin number of one equates to the border that distinguishes novel space (negative values) in MESS maps.

We illustrate the usefulness of these two methods for black grouse (Tetrao tetrix) in Switzerland (Zurell et al., 2011; for more details see Appendix S4). Conceptually, the problem is slightly different to that of the simulated species. Clearly, we do not know the true niche of the species. But we know the environmental space covered by the sample and could suppose that for predictions to other times or places, there may be combinations of environments not present in the training data. Hence, we are interested in how the model predicts to such new combinations outside the training data space (as we were for the simulated species). Again, we used a GAM with cubic smoothing splines, four degrees of freedom and no interactions and BRT with tree complexity of 1 to estimate the speciesenvironment relationship. We included six environmental predictors that covered large gradients, yet only portions of all possible combinations were present (Fig. S4). In consequence, GAM and BRT exhibited distinctly different extrapolation behaviour in the unsampled parts of the multivariate environmental space, particularly in those parts with high temperatures. These differences were not evident in conventional response plots plotted on the scale of the response, but were nicely represented by inflated response curves (Fig. 3;

Figs S5 & S6). We see the advantages of the inflated curves as: (1) they are explicit about the shape of the response at different values of other variables. While in additive models this might be deduced, especially if partial plots are fitted on the scale of the link function, it requires some careful thought and is much more apparent with our methods, especially in the case of truncated responses; (2) they make clear the responses if interactions are included in the models. The increasing popularity of methods that can optionally fit interactions if detected in the data (e.g. tree-based methods), of ensembles that might include such models and of all subsets regression where interactions are potentially allowed means that model structure might not be superficially apparent. We believe that this increasing complexity of model structure requires tools that allow exploration and understanding. Here, we believe that black grouse response fitted by GAM is more plausible than that fitted by BRT. From an ecological perspective, it seems more intuitive to assume that species response to a bioclimatic variable such as mean annual temperature gradually decreases towards physiological limits (Thuiller et al., 2004).

However, different extrapolation behaviour will only constitute a problem to model transferability if models are used to extrapolate to places with non-analogue environments in which currently unobserved portions of the environmental niche become available for prediction (Williams & Jackson, 2007; Fitzpatrick & Hargrove, 2009; Dobrowski et al., 2011). We demonstrate in Fig. S7 that plotting fitted values along each variable and comparing those obtained for training and prediction data can provide useful insights. Mapping these predictions and using environmental overlap masks to explicitly show predictions in sampled and non-analogue environmental spaces emphasizes where differences in predictions are because of extrapolation behaviour of the models. Figure 4 shows the mapped predictions of Swiss black grouse occurrence probability from GAM and BRT models. While predictions for the current environment are similar for GAM and BRT (year 2001; Fig. 4a,e), the mapped predictions for the year 2100 under climate change differ substantially (Fig. 4b,f). Using environmental overlap masks (with default number of five bins per environmental variable), we can distinguish between predictions in geographic space that are within the sampled environmental space (Fig. 4c,g) where the model is, in fact, interpolating and predictions to novel environmental space (i.e. to environmental conditions beyond the sampled ranges of the variables as in MESS maps, and to novel combinations of environmental variables; Fig. 4d,h) where the model is, in fact, extrapolating. For our Swiss black grouse example, we see that main differences between GAM and BRT predictions for the scenario of climate change indeed occur in those parts of the geographic space that exhibit novel environmental conditions compared to the sample space.

We do not intend these results as general advice about SDM algorithms. GAMs will not always extrapolate well (e.g. Elith *et al.*, 2010), and BRTs might fit responses that extrapolate in ecologically realistic ways. The important issue is that using SDMs to predict to unsampled parts of the environmental space



Figure 3 Swiss black grouse example. (a) Shows the partial dependence of back grouse occurrence to mean annual temperature for generalized additive model (GAM) and boosted regression tree (BRT). (b–c) Show the respective inflated response curves. Light grey lines and dotted dark grey lines depict the temperature effects over the full range of the other predictors (minimum, maximum, median, mean and quartiles). Light grey lines indicate combinations of environmental predictors that were observed in the sample space, while dotted dark grey lines indicate extrapolations to novel, unsampled combinations. The plots represent n = 150 Latin hypercube samples from all possible combinations of environmental predictors.



Figure 4 Predicted distributions of black grouse in Switzerland made by generalized additive model (GAM) (a–d) and boosted regression tree (BRT) (e–h). (a, e) Show current predictions, while (b, f) show predictions to climate change scenario for the year 2100 (A1F1 scenario from HadCM3 with average temperature increase of 7.7 $^{\circ}$ C and average precipitation increase of 48 mm for Switzerland). Predicted distributions are logistic outputs from low (grey) to high values (black). Environmental overlap masks (with default number of five bins per environmental variable) are used for highlighting predictions to sampled (c) + (g) and to novel environmental space (d) + (h).

is inherently risky, and uncertainty in models as well as in predictions and maps needs to be carefully assessed (Rocchini et al., 2011). The plots and maps presented here were useful for visualizing the environmental space in more than one dimension and for understanding the predicted responses in this space. Plausibility of SDM fits needs to be judged individually for any species modelled and should comply with ecological theory and prior knowledge on the species (Guisan & Thuiller, 2005; Austin, 2007). As environmental variables generally correlate, linearly and nonlinearly, we will rarely find all possible combinations in any one region (or the world). Also, species may be precluded from portions of their fundamental niche because of dispersal limitations, disturbance or biotic interactions (Colwell & Rangel, 2009). In invasive species research, it has also been demonstrated that the realized niche in the native and invaded range may differ (Broennimann & Guisan, 2008). Extrapolation behaviour may be improved by model smoothing (Elith et al., 2010) or by forcing the predicted probabilities to gradually approach zero outside observed environment (Thuiller et al., 2004). More research on the effect of including interactions in models used for extrapolation is needed; it may complicate extrapolation, and alternate means of representing the ecological response (e.g. by careful construction of predictors) might be preferable.

SUMMARY

Species distribution models would yield reliable predictions under environmental change, if the entire niche was encompassed by data, meaning that samples exist for all environmental conditions the species can occur in. However, truncated or edge niches are probably common, as not all possible environmental combinations are currently present. This may lead to erroneous predictions when extrapolating to novel environments, depending on how the model extrapolates. Thus, whenever prediction is the aim, we need to rule out unrealistic extrapolation behaviour of our models or at the very least indicate where extrapolation has occurred. The tools we provide here help to explore cases that were previously difficult to visualize.

ACKNOWLEDGEMENTS

D.Z. acknowledges partial financial support by the University of Potsdam Graduate Initiative on Ecological Modelling UPGradE. J.E. was supported by Australian Research Council grant FT0991640. B.S. was supported by the German Research Foundation (grant no. SCHR 1000/3-1 and SCHR 1000/4-2). We acknowledge the Swiss Ornithological Institute for species data provisioning and the Swiss Federal Research Institute WSL for climate data provisioning and downscaling.

REFERENCES

- Albert, C.H., Yoccoz, N.G., Thomas, C., Edwards, J., Graham, C.H., Zimmermann, N.E. & Thuiller, W. (2010) Sampling in ecology and evolution – bridging the gap between theory and practice. *Ecography*, **33**, 1028–1037.
- Austin, M.P. (2007) Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological Modelling*, **200**, 1–19.
- Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, 4, 585–589.
- Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, 10, 701–709.
- Buisson, L., Thuiller, W., Casajus, N., Lek, S. & Grenouillet, G. (2010) Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, 16, 1145–1157.

- Carnell, R. (2009) *lhs: latin hypercube samples. R package version 0.5.* Available at: http://cran.r-project.org(Comprehensive R Archive Network) (accessed 29 November 2011).
- Colwell, R.K. & Rangel, T.F. (2009) Hutchinson's duality: the once and future niche. *Proceedings of the National Academy of Sciences United States of America*, **106**(Suppl. 2), 19651–19658.
- Dobrowski, S.Z., Thorne, J.H., Greenberg, J.A., Safford, H.D., Mynsberge, A.R., Crimmins, S.M. & Swanson, A.K. (2011)
 Modeling plant ranges over 75 years of climate change in California, USA: temporal transferability and species traits. *Ecological Monographs*, 81, 241–257.
- Elith, J. & 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. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677–697.
- Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distribution from occurrence data. *Ecography*, **29**, 129–151.
- Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.
- Fitzpatrick, M.C. & Hargrove, W.W. (2009) The projection of species distribution models and the problem of non-analog climate. *Biodiversity Conservation*, **18**, 2255–2261.
- Franklin, J. (2010) Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, **16**, 321–330.
- 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.
- Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkkala, R., Thuiller, W. & Sykes, M.T. (2006) Methods and uncertainties in bioclimatic envelope modelling under climate change. *Pro*gress in Physical Geography, **30**, 1–27.
- Medley, K.A. (2010) Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus* Skuse (Culicidae), revealed by reciprocal distribution models. *Global Ecology and Biogeography*, **19**, 122–133.
- Pearson, R.G., Thuiller, W., Araújo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T.P. & Lees, D.C. (2006) Model-based uncertainty in species range prediction. *Journal of Biogeography*, 33, 1704– 1711.
- Platts, P.J., McClean, C.J., Lovett, J.C. & Marchant, R. (2008) Predicting tree distributions in an East African biodiversity hotspot: model selection, data bias and envelope uncertainty. *Ecological Modelling*, 218, 121–134.

- R Development Core Team. (2010) *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.
- Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jiménez-Valverde, A., Ricotta, C., Bacaro, G. & Chiarucci, A. (2011) Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Progress in Physical Geography*, **35**, 211–226.
- Thuiller, W. (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020–2027.
- Thuiller, W., Brotons, L., Araújo, M.B. & Lavorel, S. (2004) Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, **27**, 165– 172.
- Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 868–883.
- Williams, J.W. & Jackson, S.T. (2007) Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, **5**, 475–482.
- Williams, J.W., Jackson, S.T. & Kutzbach, J.E. (2007) Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Sciences United States* of America, **104**, 5738–5742.
- Zurell, D., Berger, U., Cabral, J.S., Jeltsch, F., Meynard, C.N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B. & Grimm, V. (2010) The virtual ecologist approach: simulating data and observers. *Oikos*, **119**, 622– 635.
- Zurell, D., Grimm, V., Rossmanith, E., Zbinden, N., Zimmermann, N.E. & Schröder, B. (2011) Uncertainty in predictions of range dynamics: black grouse climbing the Swiss Alps. *Ecography*, doi: 10.1111/j.1600-0587.2011.07200.x.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1 Response surfaces for simulated species 1.

Figure S2 Fitted values for simulated species 2.

Figure S3 Response surfaces for simulated species 3.

Figure S4 Environmental scatterplots for black grouse occurrences.

Figure S5 Inflated response curves of GAM for black grouse occurrence.

Figure S6 Inflated response curves of BRT for black grouse occurrence.

Figure S7 Fitted values for black grouse occurrence.

Appendix S1 Code for simulated data and visualisation tools.

D. Zurell et al.

Appendix S2 Supporting information for simulated data example.

Appendix S3 Method descriptions for visualisation tools.

Appendix S4 Supporting information for Swiss black grouse example.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCHES

Damaris Zurell is a research fellow at the University of Potsdam with special interests in landscape ecology, conservation biogeography and behavioural ecology. Her research combines statistical and mechanistic modelling approaches to study how environmental change affects broad-scale species' distribution, population dynamics as well as individual fitness. This contribution was part of her PhD thesis. **Jane Elith** is an Australian Research Council Future Fellow based at the University of Melbourne. She specializes in methods for implementing and evaluating species distribution models with a focus on relevance to intended applications. Her current projects span terrestrial, freshwater and marine ecosystems and include invasive species and climate change applications.

Boris Schröder is Professor for Landscape Ecology at the Technische Universität München. His research is dedicated to understanding the relationship between patterns, processes and functions in dynamic landscapes as well as the development of models for the conservation and sustainable management of species, landscapes and related ecosystem functions and services.

Editor: Janet Franklin