

DISENTANGLING INTERPOLATION AND EXTRAPOLATION UNCERTAINTIES IN ECOLOGICAL NICHE MODELS: A NOVEL VISUALIZATION TECHNIQUE FOR THE SPATIAL VARIATION OF PREDICTOR VARIABLE COLINEARITY

JAN O. ENGLER AND DENNIS RÖDDER

Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, D - 53113 Bonn, Germany

Abstract. – Environmental niche models (ENMs) are increasingly used in many scientific fields, with most studies requiring the application of the ENM to predict the likelihood of occurrence and/or environmental suitability in locations and time periods outside the range of the data set used to fit the model. Uncertainty in the quality of ENM predictions caused by errors of interpolation and extrapolation has been acknowledged for a long time, but the explicit consideration of the magnitude of such errors is, as yet, uncommon. Among other issues, the spatial variation in the colinearity of the environmental predictor variables used in the development of ENMs may cause misleading predictions when applying ENMs to novel locations and time periods. In this paper, we provide a framework for the spatially explicit identification of areas prone to errors caused by changes in the inter-correlation structure (i.e. their colinearity) of environmental predictors used for ENM development. The proposed method is compatible with all ENM algorithms currently employed, and expands the available toolbox for assessing the uncertainties rising from ENM predictions. We provide an implementation of the analysis as a script for the R statistical platform in an online appendix.

Key words. – climate change, environmental niche, residual distribution, *Plethodon*, PURV plots, R statistical platform

INTRODUCTION

The development in computational resources during the last few decades, combined with an increasing availability of environmental data and information on species occurrences, has led to a boost in the application of environmental niche models (ENMs) in many biological fields (Kozak *et al.* 2008, Elith and Leathwick 2009). In correlative ENMs, an idealized environmental niche of the target species is derived from the environmental conditions found at the locations where the species is known to occur and/or absence or pseudo-absence data reflecting the general conditions within this area (Guisan and Zimmermann 2000, Peterson *et al.* 2011). Analyzing the environmental conditions at the species' occurrence records, ENMs can be used to estimate the target species' potential distribution based on its realized niche, which is commonly a subset of its fundamental niche (Soberón and Peterson 2005, Peterson *et al.* 2011). Based on the ENM, the derived habitat preferences can be subsequently projected into unsampled locations and time periods using a geographic information system (GIS).

Various methods have been proposed for this purpose (Guisan and Zimmermann 2000, Elith and Leathwick 2009, Franklin 2009,

Peterson *et al.* 2011). Early ENM approaches define the species' niche as a multidimensional boxcar or convex hull envelope enclosing the environmental characteristics of the known occurrences of the species in environmental space (e.g., BIOCLIM, Busby 1991, DOMAIN, Carpenter *et al.* 1993). However, more recently developed methods, such as artificial neural networks (Olden *et al.* 2008), multivariate adaptive regression splines (Friedman 1991), random forest (Breiman 2001), and maximum entropy approaches (Phillips *et al.* 2006), allow for the incorporation of even more complex interactions between predictor variables. These techniques commonly require absence or pseudo-absence data for model training. This extra flexibility has been shown to increase the accuracy of predictions computed with these methods, frequently outperforming more conventional approaches (Elith *et al.* 2006, Hernandez *et al.* 2006, Wisz *et al.* 2008). However, appropriate selection of absence or pseudo-absence data requires special attention here since the choice may strongly influence the reliability and interpretation of the results (Saupe *et al.* 2012). Various authors have suggested that the most appropriate background data should reflect the environmental space that is potentially

colonizable by the target species (e.g., Anderson and Raza 2010, Barve *et al.* 2011, Saupe *et al.* 2012). Although this restriction of the training area of an ENM may result in a more specific model, it is at the same time commonly less generalizable across space and time.

Applications of ENMs are manifold, ranging, for example, from simple visualizations of species' potential distributions (e.g., Brown and Twomey 2009), assessments of the potential distribution of invasive species (e.g., Peterson and Vieglais 2001, Phillips *et al.* 2008), quantification of possible impacts of climate change (e.g., Araújo *et al.* 2004, Thomas *et al.* 2004), to analyses of niche evolution and speciation (e.g., Kozak and Wiens 2006, Peterson 2011). Many of these applications require some degree of model prediction to novel locations and time periods. However, with increasing model complexity, assessing the quality of the prediction resulting from such a transfer is becoming more difficult (Randin *et al.* 2006, Peterson 2007, Elith *et al.* 2010, Rödder and Lötters 2010, Peterson *et al.* 2011), and various possible error sources related to issues of methodology and the biological characteristics of the species have been identified (Heikkinen *et al.* 2006, Franklin 2009, Jiménez-Valverde *et al.* 2009, Dobrowski *et al.* 2011, McNerny and Purves 2011, Rocchini *et al.* 2011). Among these, changes in the predictor colinearity through space and time may cause errors, not only when extrapolating, but also when interpolating ENMs within the area of interest (Elith and Leathwick 2009, Jiménez-Valverde *et al.* 2009, Elith *et al.* 2010, Rocchini *et al.* 2011). Moreover, it has been suggested that some of the more sophisticated methods that use both information found from areas where the species is present and from where it is absent (presence/absence models) may characterize a species' realized niche, where others, which only use the information contained in the records where species are known to occur (presence/pseudo-absence or presence-only models) quantify its potential distribution (Rödder and Lötters 2010, Jiménez-Valverde *et al.* 2011, Peterson *et al.* 2011). These conceptual differences require additional attention when interpreting ENM predictions and different types of uncertainties may arise when applying one over the other method.

As stated above, the selection of appropriate predictor variables is a critical task during ENM development and is one that is likely to dramatically influence the modeling results (e.g.,

Peterson and Nakazawa 2008, Rödder *et al.* 2009, Synes and Osborne 2011, Peterson *et al.* 2011, Varela *et al.* 2011). Generally, environmental predictors used in ENMs can be classified in a continuum spanning between proximal or distal predictors, depending on how they affect the fitness of the target species (Austin 2002). Proximal predictors are those actually affecting the physiology of the species, whereas distal variables do not directly affect the target species. Unfortunately, it is commonly not possible to use the most proximal variables due to limited availability. Hence, for most ENM approaches, predictors are used that are rather intermediate or even distal within the continuum. Nonetheless, within a complete set of imaginable environmental predictors, they may still be correlated with the most proximal predictors and therefore be informative. When assessing the correlation structure among all possible combinations of available predictors, this information can be used to estimate the correlation of the available predictors with those (unavailable) proximal predictors.

Distal predictors should only be used for ENM development if a high degree of colinearity with proximal predictors exists – an assumption that is frequently fulfilled. Modern ENM algorithms search for a set of predictors or derived features thereof that best explain the variation in the dependent variable: the presence and/or absence of the target species across the different environments. As in a standard multiple regression model, these methods can supplement the supplied environmental predictors with extra predictor variables representing interactions between the predictors. Within the region an ENM is trained for, effects of proximal predictors can be estimated by distal predictors as long as both are correlated.

These relationships and the implications for the transfer of the model to novel geographic regions are exemplified in Figure 1: consider the high degree of correlation between the minimum temperature of the coldest month (assumed proximal variable) and the maximum temperature of the warmest month (assumed distal variable) within the 'training area' and its deviance in other areas. When transferring an ENM onto environments differing from the training conditions, predictions may only be useful in those cases where the predictor colinearity between interacting variables in the ENM as well as between proximal and distal variables remain stable. In those areas where proximal

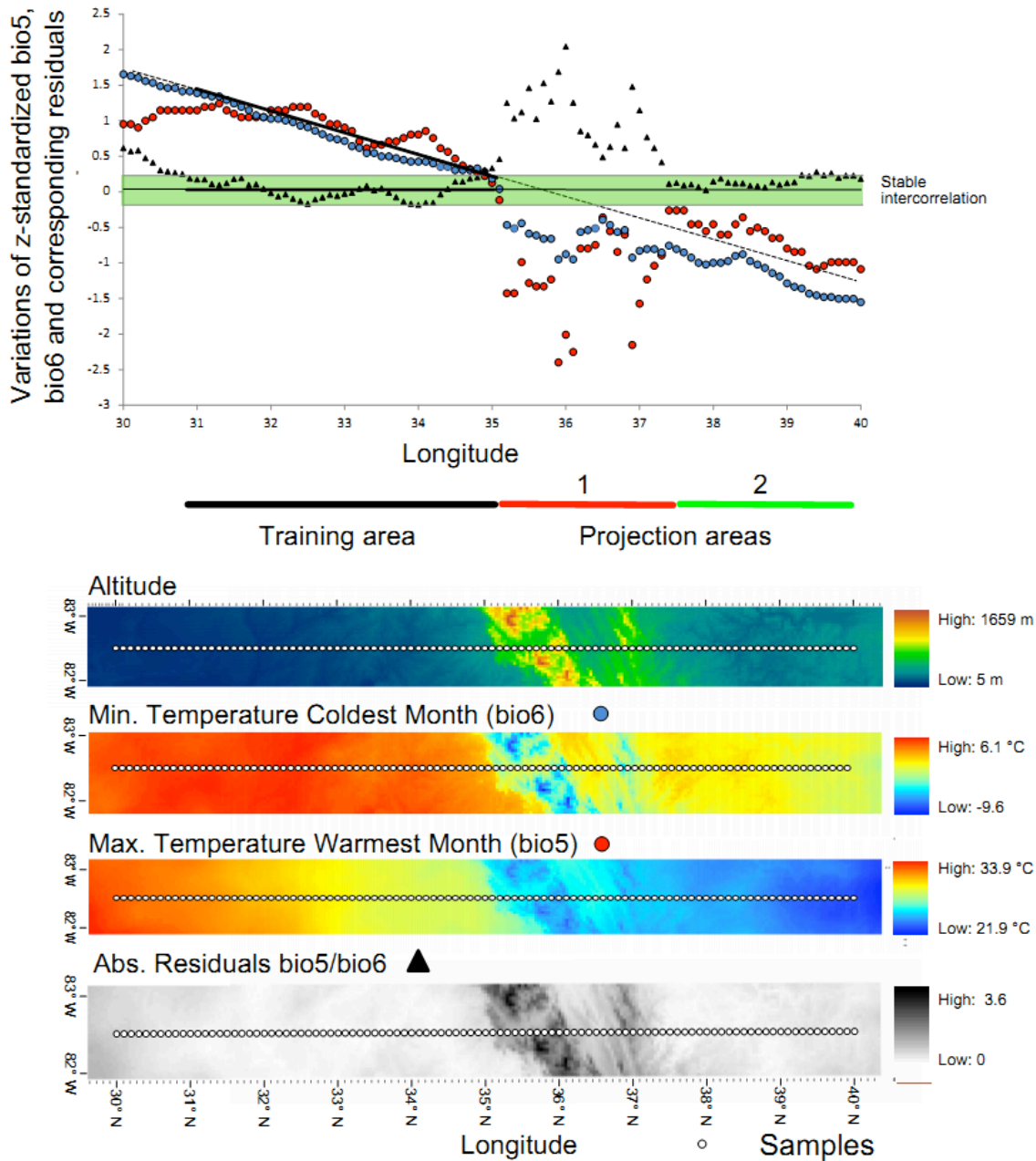


Figure 1: Illustration of the potential effects of local changes in the inter-correlation between two predictors across a longitudinal gradient in North America introduced by topographic factors as well as the underlying principles of PURV (*Prediction Uncertainty Assessments using Residual Variation*). Assume that an ENM is developed within one part of the study region (training area, black bar) and subsequently projected onto the rest of the gradient. The variability of the residuals within the training region can be used as reference to quantify the degree of inter-correlation between both predictors as incorporated in the ENM. When projecting the ENM through space or time, any change in the amplitude of residuals exceeding the variation present in the training area may indicate a change in the local inter-correlation, e.g. as present in projection area 1. Projections outside of the training area of an ENM may only be reliable when the amplitude of the residuals is not larger than in the training region (e.g. in projection area 2). The magnitude of the deviance can be quantified using the proportion of the residuals at a given grid cell exceeding the confidence limits within the training region.

physiological constraints affecting a species' survival may not be well reflected in the distal variables, model projections may become unreliable. In Figure 1, we see that this is the case in 'projection area 1' where any prediction of an

ENM developed in the 'training area' may be misleading. In contrast, the degree of inter-correlation between both predictors is well within the training range in projection area 2 and, as a result, predictions may be much more reliable.

Aside from possible uncertainties arising from the inter-correlation structure between proximal and distal variables, the inclusion of highly correlated variables may violate model assumptions in many regression techniques. Therefore, *a priori* selection of a set of less correlated variables (e.g., $R^2 < 0.75$) is often necessary, wherein the putatively biologically most relevant predictors should be selected (e.g., Saupe *et al.* 2012). As long as the inter-correlation structure among selected and omitted predictors is stable, this information reduction is appropriate – but not when it changes, since in this case ENM projections may become unreliable.

Although several authors mention likely problems caused by varying inter-correlation structures of predictors between training and projection conditions (e.g., Heikkinen *et al.* 2006, Peterson 2007, Jiménez-Valverde *et al.* 2009, Elith *et al.* 2010), no technique allowing for a spatially-explicit evaluation of this error source is available, although Elith *et al.* (2010) presented some general ideas. Here, we propose the use of *Prediction Uncertainty Assessments using Residual Variation* (PURV) plots based on comparisons between the residual ranges of each pair of predictors within the training area of an ENM and the projection areas to produce maps showing spatial explicit variations in correlation matrices. These plots can be used to identify areas where ENM predictions may be prone to errors, which may lead to erroneous conclusions when interpreting prediction maps.

METHODS

Species Distribution Models

In order to illustrate the applicability of PURV plots to assess variations in correlation structures, we developed ENMs for two sister species of North American salamanders (i.e. *Plethodon cylindraceus* (Harlan, 1825) and *P. teyahalee* Hairston, 1950) as described in Kozak and Wiens (2006). A total of 180 georeferenced species records of *P. cylindraceus* and 761 records of *P. teyahalee* were obtained through the Global Biodiversity Information Facility (accessed through GBIF data portal¹; NMNH Vertebrate Zoology Herpetology Collections²; MVZ Herp Catalog³; Herp Specimens⁴) and checked for possible georeferencing errors using

DIVA-GIS 7.1.6 (Hijmans *et al.* 1999, Hijmans *et al.* 2005a).

Information on current climate was extracted from the Worldclim database (Hijmans *et al.* 2005b⁵). The set of predictor variables used for ENM development consisted of those variables identified by Kozak and Wiens (2006) as being biologically relevant to the salamanders: ‘mean annual temperature range’, ‘maximum temperature of the warmest month’, ‘minimum temperature of the coldest month’, ‘precipitation seasonality’, and ‘precipitation of the driest quarter’.

Using Maxent 3.3.3e (Phillips *et al.* 2006, Phillips and Dudík 2008, Elith *et al.* 2010) we developed ENMs for both species. Maxent requires the designation of a set of pseudo-absences for the calculation of species habitat preferences and, to this end, random background points were automatically sampled by Maxent in a circular buffer around each record of 0.16° (~ 14 km radius), ensuring that each grid cell was considered only once to avoid pseudo-replication. This area reflects the geographic space potentially accessible for the species (Phillips *et al.* 2009, Anderson and Raza 2010, Barve *et al.* 2011, Saupe *et al.* 2012). To assess the performance of the model, 100 ENMs per species were developed using the default Maxent settings, but splitting the species records into 70 % used for training the model and 30 % to test model performances by calculating the Area Under the receiver operating Curve (AUC) (Hanley and McNeil 1982, Swets 1988). Choosing the logistic output format with values ranging linearly from 0 (unsuitable) to 1 (optimal) (Phillips and Dudík 2008) the averages and standard deviations of the potential distributions of the species suggested by the 100 ENMs were projected onto a larger geographic area. The resulting maps show each species’ probability of occurrence per grid cell and its variation across all replicates.

The degree of environmental novelty when extrapolating the ENMs into conditions outside those found in the training region was assessed using the application of multidimensional environmental similarity surfaces (MESS) as described by Elith *et al.* (2010) (i.e., using the relevant tool implemented in Maxent 3.3.3e). In MESS maps, values range from +100 to $-\infty$, with positive values indicating grid cells with environmental conditions within the range of the

¹ www.gbif.org

² <http://data.gbif.org/datasets/resource/1838>

³ <http://data.gbif.org/datasets/resource/8123>

⁴ <http://data.gbif.org/datasets/resource/8956>

⁵ www.worldclim.org

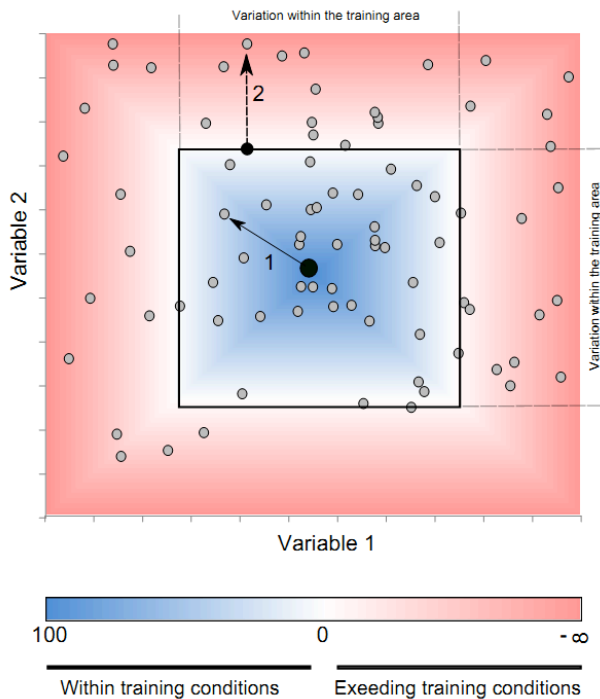


Figure 2: Principles of Multivariate Environmental Similarity Surfaces (MESS; Elith *et al.* 2010): if the observed value p_i at grid cell P_i lies within the range of Variable V_i in the training region, the similarity of this grid cell with respect to V_i is defined as percentage deviance from the median of the observed range (arrow 1). This creates a gradient across the environmental envelope within the training area ranging from 100 in the center to 0 at the margins indicated by the black box. In multivariate environmental space this approach is equal to the BIOCLIM approach (Busby 1991). In those cases where any p is outside of this environmental envelope spanned by V within the training region, the distance to its margin of the most divergent Variable V_i is measured (arrow 2), wherein the scores reflect the quotient of the absolute deviance of p_i from the margin of the envelope's range. A score of -100 would indicate a deviance of p_i equaling the range of V_i within the training area. In multivariate space, the score of the most divergent V_i at P_i is assigned. In PURV plots, grids showing residuals are used instead of environmental variables themselves.

variables in the training region (analogous to the cells predicted as suitable in the BIOCLIM algorithm; Busby 1991) and negative values denoting grid cells with environmental conditions that fall outside the range of environments present in the training area (as described by the boxcar environmental envelope – see Figure 2). Note that the training region of an ENM ideally comprises those areas that are potentially colonizable for the target species, wherein a number of different configurations are possible (Saupe *et al.* 2012). MESS maps were transformed into binary

shapefiles, indicating those areas where at least one variable falls outside the range present in the training area of the ENM.

Prediction Uncertainty Assessments using Residual Variation (PURV)

For explicit spatial comparisons of correlation structures among predictors across geographic space, we developed a function for the R statistical platform for the construction of PURV plots (Supplementary material Appendix 1). As reference for subsequent analyses, we assessed the inter-correlation structure of each pair of z-standardized predictor variables within the training area of the ENM. Z-standardization referred to the average and standard deviation of the respective variable within the training area. Therefore, we used simple linear regressions and determined the intercept and slope of the corresponding regression. We used linear instead of other, more complex, relationships as basic assumption for this analysis, as the true inter-correlation structure for a given projection area is *a priori* not known and might change between different predictor variable comparisons. Therefore, assuming a linear relationship represents the most conservative approach for the construction of PURV plots. For each pair of predictor variables, we created a grid covering the previously-defined projection area showing the spatial distribution of residuals in R 2.12 (R Development Core Team 2011), which is parameterized based on the correlation structure in the training area only. These grids allow for a spatially-explicit quantification of the variation in the correlation structure within both the training and projection areas, where the magnitude of the variation within the residuals in the training area of the ENM can be used as reference. In those projection areas where the residuals exceed the range observed within the training area, it can be assumed that the two variables are less correlated locally.

The number of possible pair-wise comparisons of predictors can quickly become large, making detailed analyses time consuming. Therefore, MESS analyses can be used to summarize the maximum possible effect of inter-correlation changes as described by Elith *et al.* (2010). The novelty here is that the MESS plots are derived from the residual grids instead of the original environmental predictors. Wherein traditional MESS analyses ask whether the environmental conditions at a given site are within those conditions available within the

training area, our PURV approach asks whether the inter-correlation structure of the environmental variables at a given site are similar to those within the training area or not. PURV plots allow the quantification of errors in two different areas (compare Figure 2): (1) in regions characterized by the same inter-correlation structure present in the training region, the plots show the relative distance of each grid cell to the center of the multivariate boxcar residual envelope, and (2) in areas in which the inter-correlation structure deviates from the inter-correlation structure within the training area, the plots show the distance to the border of the boxcar envelope of the most distant residual. In the former case positive values ranging from 100 (center of the boxcar envelope) to 0 (its margin) are assigned, while in the latter case, negative scores are assigned, identifying the maximum deviant change in the predictor correlation (for a more detailed description of the MESS procedure see Elith *et al.* 2010).

ENM projections into areas where the range of residuals in at least one predictor-pair is greater than that found within the training area can be interpreted as being less reliable. This becomes reasonable when considering two different cases: (1) interactions between two variables fitted by an ENM refer to the correlation structure of the predictors within the training area. Unreasonable response curves might result when projecting these relationships onto deviating correlation structures in other areas or time slices. (2) In many cases it is necessary to select only the putatively biologically most important variable for ENM development if a set of variables is highly correlated. As long as the correlation structure remains stable this subjective selection may cause no negative effects, but in those cases where the correlation structure between a selected variable and one omitted variable changes, the selected variable cannot be used to estimate the omitted ones. The PURV plots indicate spatial changes in the correlation structure among predictor variables irrespective and unaffected from the ENM algorithm. For ease of interpretation, it may be helpful to transform the continuous scale of the residuals deviation into a binary classifier where a '1' value is given to any cell with a residual deviation found within the confidence interval of the residual deviation of cells within the training region. All other cells would be given a value of '0'.

RESULTS

A summary of the model performances in terms of training and test AUCs, relative contributions of the environmental predictors to the final ENM, and presence/absence thresholds is given in Table 1. AUC scores on average indicate good model performance and ENMs developed for both *Plethodon cylindraceus* and *P. teyahalee* depict the known distributions of the species very well. However, they also project high probabilities of occurrence in some areas outside the known range of each salamander species (Figure 3a, b), many of which are situated in the known range of the other species. This is not unexpected since both are sister taxa and likely occupy similar environmental niches (Kozak and Wiens 2006), thus a mixture of dispersal limitation and biotic interaction most likely explains the lack of this species in areas deemed suitable. As indicated by the SD per grid cell, some of the projected potentially suitable areas for each species show a comparatively high variability ($SD_{\max} = 0.29$ in both) (Figure 3, middle row) suggesting a higher degree of model uncertainty. Only MESS analyses based on the original predictor variables highlight parts of them as prone to potential errors (Figure 3, bottom). Moreover, the areas identified by PURV plots only partly overlap with them, indicating that the spatial distribution of both error sources is not necessarily coincident. This appears to be reasonable since the extrapolation situation identified by MESS refers to the range of the variables within the training region, but the PURV plots identify changes in the correlation structure among predictors. Even if the range of predictors exceeds the environmental conditions within the training area as indicated by MESS, the correlation structure might still be within the training range. Vice versa, this might also explain the spatial independence of MESS and PURV plots.

DISCUSSION

Our results highlight the importance of considering both the relative ranges of environmental parameters as well as their inter-correlation structures in the training region when projecting ENMs into novel locations and time periods.

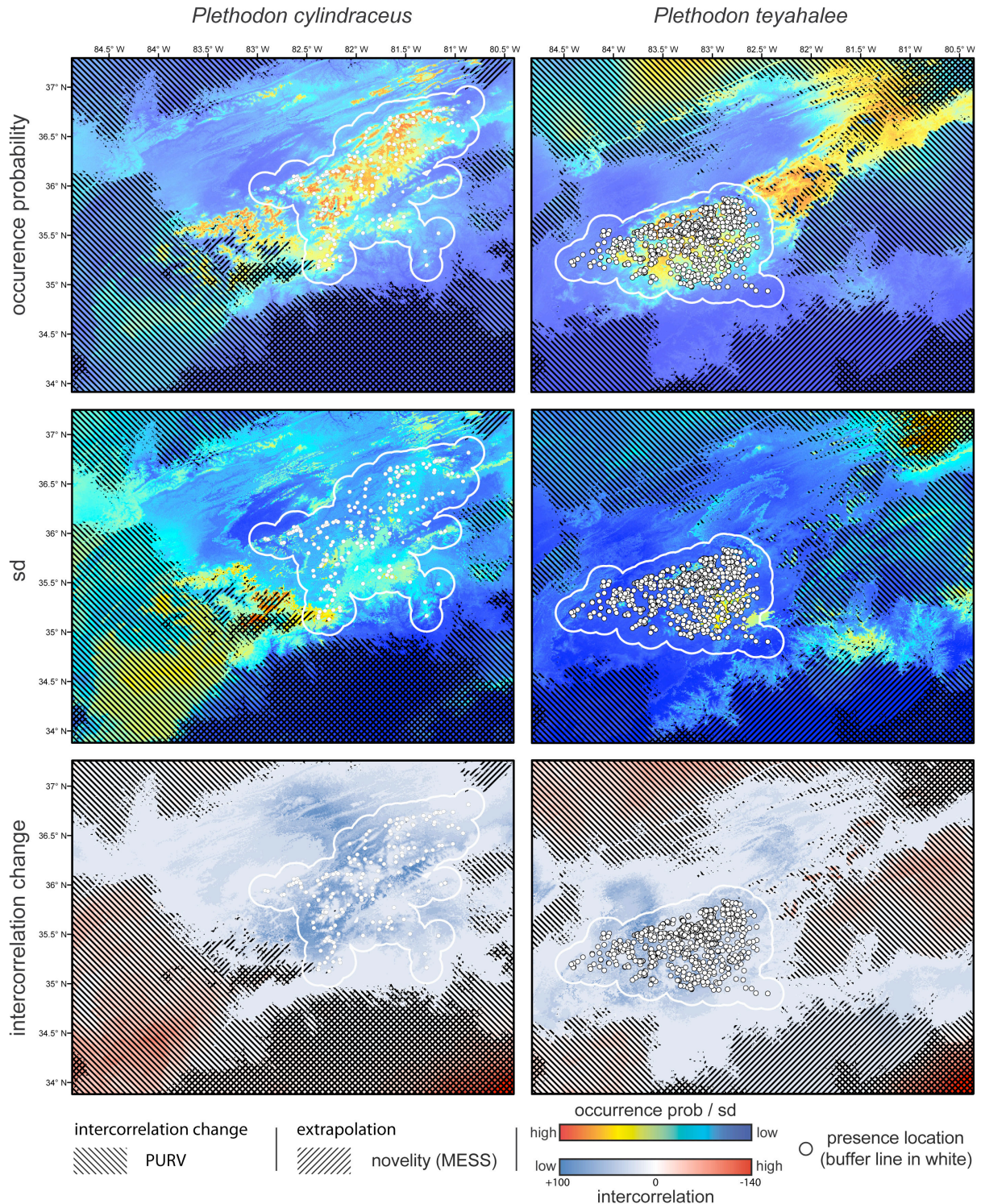


Figure 3: Occurrence probability maps of *Plethodon cylindraceus* (a) and *P. teyahalee* (b), corresponding maps of standard deviations of 100 Maxent models (c, d) and *Prediction Uncertainty assessment using Residual Variation* analysis (PURV) plots (e, f). Regions highlighted by PURV confidence limits are indicated as downward hatched, wherein areas requiring model extrapolation identified by Multidimensional Environmental Similarity Analyses (MESS) are upward hatched. Models were trained with species records (white dots) and random background points drawn within circular buffers (white outlines).

Extrapolation errors

In both species, multivariate environmental similarity surfaces (MESS) based on the original environmental predictors identified varying

proportions of the studied areas in which at least one variable exceeded its range present in the training region (Figure 3, Appendix A). Here, assigning any occurrence probability requires

extrapolation by the model. In most cases, the ENMs did not predict the species in those novel environments, but where they did, SD of the corresponding occurrence probabilities became much higher compared to other regions, thus depicting extrapolation areas quite well (e.g. in central areas in *P. cylindraceus* and in the north-eastern corner of *P. teyahalee*). Predictions in these areas are much more variable than interpolated predictions due to the requirement of model extrapolation, and hence less reliable. Other portions of the study area where the species were not predicted to occur but requiring extrapolation such as the south-eastern areas were correctly identified as being too hot in the warmest month in *P. cylindraceus* or too cold in the coldest month in *P. teyahalee*. This was correctly incorporated in the ENMs by fading response curves at the corresponding tail of the temperature range (Appendix A). Note that both predictors, the maximum temperature of the warmest month and the minimum temperature of the coldest month, had the highest overall variable contribution in the respective ENMs (Table 1). Uncertainties due to extrapolation may therefore only arise when high response values for a given parameter are extrapolated but not when physiological limits are correctly captured by fading response curves. For further discussion how the shape of response curves may affect the reliability of ENMs, see Santika and Hutchinson (2009).

Spatial changes in the inter-correlation structure

PURV plots indicated that the correlation structure of the variables in the training area was only partly coincident with those in the projection area (Figure 3). A number of spots in the study areas were identified as showing substantial differences in their correlation structures. However, the species were only occasionally predicted to occur in the areas highlighted by PURV plots (e.g. northern and north-eastern parts of the study area in *P. teyahalee*). Unlike in areas requiring extrapolation of the ENM beyond the parameter range within the training region, no relationship between SD in the ENMs and inter-correlation change was evident, suggesting that this kind of error source may frequently remain undetected in traditional analyses.

Jiménez-Valverde *et al.* (2009) identified fairly-to-highly similar correlation structures on a continental scale in the most commonly used climate data sets using Mantel tests, where the correlation structures were most similar when comparing different time slices than when comparing different areas. Whilst these comparisons on a continental scale may indicate detectable but rather low variations, this may not necessarily be true on smaller scales. We obtained quite different PURV plots for both *Plethodon* species, suggesting that spatial position of the areas exhibiting the largest

Table 1. Performance and characteristics of Maxent models for *Plethodon cylindraceus* and *Plethodon teyahalee*. AUC = area under the receiver operating characteristic curve; Min train = minimum training presence; 10 % train = lowest 10 percentile training omission.

Species	<i>P. cylindraceus</i>	<i>P. teyahalee</i>
<i>Model performance</i>		
Training AUC	0.836	0.802
Test AUC	0.807	0.790
<i>Variable contributions [%]</i>		
mean annual temperature range	17.4	6.8
maximum temperature warmest month	57.9	22.4
minimum temperature coldest month	5.3	61.9
precipitation seasonality	2.5	3.5
precipitation driest quarter	16.9	5.3
<i>Presence/absence thresholds</i>		
Min training	0.065	0.034
10 % training	0.239	0.252

changes in the correlation structure strongly depend on the position of the training area. On this small scale, the spatial variability of the residuals within the training areas can be much smaller than can be expected on a continental scale due to fine-scale topographic features. This increases the probability of locally-occurring high inter-correlations among variables resulting in a higher chance of deviations across space - even within a small area. As a consequence, although the negative impacts caused by changes in the inter-correlation structure of predictors may be rather small on large scales, they may bear heavily on ENM predictions for species occupying rather small ranges (e.g. the two *Plethodontid* salamanders).

CONCLUSIONS

As illustrated by our results, PURV plots are able to highlight formerly undetected areas of uncertainty in projection areas of ENMs, which are *per se* independent of the employed algorithm. However, the impact of the uncertainty caused by changes in the correlation structure of predictors may strongly depend on the algorithm used. It may be absent when using simple profile algorithms, which do not incorporate variable interactions at all such as BIOCLIM (Busby 1991) or DOMAIN (Carpenter *et al.* 1993), where an appropriate choice of relevant proximal predictors may be more important. In contrast, the impact of changes in the correlation structure of predictors may become highest when interactions among predictors are incorporated in the model, as is commonly the case with state-of-the-art algorithms (Elith and Leathwick 2009, Franklin 2009, Thuiller *et al.* 2009), such as artificial neural networks, boosted regression trees, generalized additive models, classification and regression trees, random forests or Maxent. Furthermore, projection areas identified using PURV plots might become unreliable when inter-predictor relationships deviate from simple colinearity into more complex relationships, since residual deviation may be increased. There is a need for future extensions incorporating more complex relationships into this approach, and the authors welcome useful comments from the scientific community. Summarizing, PURV plots in its current stage may be very helpful to identify areas prone to misleading predictions, an area of research that requires more attention in the future use of ENMs.

ACKNOWLEDGMENTS

We are grateful to Kai F. Abt, Joseph Chipperfield and Ursula Bott for helpful comments on an early version of the manuscript. D.R. is grateful for financial support by the Ministry of Education, Science, Youth and Culture of the Rhineland-Palatinate state of Germany within the framework of the project 'Implications of global change for biological resources, law and standards.' Comments by A. Townsend Peterson and an anonymous reviewer greatly improved the manuscript.

REFERENCES

- Anderson, R.P., and A. Raza (2010). The effect of the extent of the study region on GIS models of species geographic distribution and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography* 37: 1378-1393.
- Araújo, M.B., M. Cabeza, W. Thuiller, L. Hannah, and P.H. Williams (2004). Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology* 10: 1618-1626.
- Austin, M.P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* 157: 101-118.
- Barve, N., V. Barve, A. Jiménez-Valverde, A. Lira-Noriega, S.P. Maher, A. T. Peterson, J. Soberón, and F. Villalobos (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling* 222: 1810-1819.
- Breiman, L. (2001). Random forests. *Machine Learning* 45: 5-32.
- Brown, J.L., and E. Twomey (2009). Complicated histories: three new species of poison frogs of the genus *Ameerega* (Anura: Dendrobatidae) from north-central Peru. *Zootaxa* 2049: 1-38.
- Busby, J.R. (1991). BIOCLIM - a bioclimatic analysis and prediction system. In: Margules, C.R. and M.P. Austin (eds.), *Nature conservation: cost effective biological surveys and data analysis*. CSIRO, Melbourne, pp. 64-68.
- Carpenter, G., A.N. Gillison, and J. Winter (1993). DOMAIN: a flexible modeling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation* 2: 667-680.
- Dobrowski, S.Z., J.H. Thorne, J.A. Greenberg, H.D. Safford, A.R. Mynsberge, S.M. Crimmins, and A.K. Swanson (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., and J.R. Leathwick (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual*

- Reviews in Ecology, Evolution and Systematics 40: 677-697.
- Elith, J., C.H. Graham, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. McC.M. Overton, A. Townsend Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberón, S. Williams, M.S. Wisz, and N.E. Zimmermann (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129-151.
- Elith, J., M. Kearney, and S. Phillips (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution* 1: 330-342.
- Franklin, J. (2009). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.
- Friedman, J. (1991). Multivariate adaptive regression splines. *Annals of Statistics* 19: 1-141.
- Guisan, A., and N. Zimmermann (2000). Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147-186.
- Hanley, J., and B. McNeil (1982). The meaning of the use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143: 29-36.
- Heikkinen, R.K., M. Luoto, M.B. Araújo, R. Virkkala, W. Thuiller, and M.T. Sykes (2006). Methods and uncertainties in bioclimatic envelope modeling under climate change. *Progress in Physical Geography* 30: 751-777.
- Hernandez, P.A., C.H. Graham, L.L. Master, and D.L. Albert (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29: 773-785.
- Hijmans, R.J., M. Schreuder, J. De la Cruz, and L. Guarino (1999). Using GIS to check coordinates of genebank accessions. *Genetic Resources and Crop Evolution* 46: 291-296.
- Hijmans, R.J., L. Guarino, A. Jarvis, R. O'Brien, P. Mathur, C. Bussink, M. Cruz, I. Barrantes, and E. Rojas (2005a). *DIVA-GIS Version 5.2 Manual*.
- Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis (2005b). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978.
- Jiménez-Valverde, A., Y. Nakazawa, A. Lira-Noriega, and A.T. Peterson (2009). Environmental correlation structure and ecological niche model projections. *Biodiversity Informatics* 6: 28-35.
- Jiménez-Valverde, A., A.T. Peterson, J. Soberón, J.M. Overton, P. Aragón, and J.M. Lobo (2011). Use of niche models in invasive species risk assessments. *Biological Invasions* 13: 2785-2797.
- Kozak, K., and J.J. Wiens (2006). Does niche conservatism promote speciation? A case study in north American salamanders. *Evolution* 60: 2604-2621.
- Kozak, K.H., C.H. Graham, and J.J. Wiens. (2008). Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology and Evolution* 23: 141-148.
- McInerney, G.J., and D.W. Purves (2011). Fine-scale environmental variation in species distribution modeling: regression dilution, latent variables and neighborly advice. *Methods in Ecology and Evolution* 2: 248-257.
- Olden, J.D., J.J. Lawler, and N.L. Poff (2008). Machine learning methods without tears: a primer for ecologists. *Quarterly Reviews in Biology* 83: 171-183.
- Peterson, A.T. (2007). Why not Why Where: The need for more complex models of simpler environmental spaces. *Ecological Modelling* 203: 527-530.
- Peterson, A.T. (2011). Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography* 38: 817-827.
- Peterson, A.T., and Y. Nakazawa (2008). Environmental data sets matter in ecological niche modeling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography* 17: 135-144.
- Peterson, A.T., and D.A. Vieglais (2001). Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. *BioScience* 51: 363-371.
- Peterson, A.T., J. Soberón, R.G. Pearson, R.P. Anderson, E. Martínez-Meyer, M. Nakamura and M.B. Araújo (2011). *Ecological Niches and Geographic Distributions*. Princeton University Press, Princeton, 328 pp.
- Phillips, B.L., J.D. Chipperfield, and M.R. Kearney (2008). The toad ahead: challenges of modelling the range and spread of an invasive species. *Wildlife Research* 35: 222-234.
- Phillips, S.J., and M. Dudík (2008). Modeling of species distributions with Maxent: new extensions and comprehensive evaluation. *Ecography* 31:161-175.
- Phillips, S.J., R.P. Anderson, and R.E. Schapire (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231-259.
- Phillips, S.J., M. Dudík, J. Elith, C.H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications* 19: 181-197.
- R Development Core Team (2011). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. R Foundation for Statistical Computing.
- Randin, C.F., T. Dirnböck, S. Dullinger, N.E. Zimmermann, M. Zappa, and A. Guisan (2006). Are niche-based species distribution models transferable in space? *Journal of Biogeography* 33: 1689-1703.

- Rocchini, D., J. Hortal, S. Lengyel, J.M. Lobo, A. Jiménez-Valverde, C. Ricotta, G. Bacaro, and A. Chiarucci (2011). Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. *Progress in Physical Geography* 35: 211-226.
- Rödder, D., and S. Lötters (2010). Explanative power of variables used in Species Distribution Modelling: An issue of general model transferability or niche shift in the invasive Greenhouse frog (*Eleutherodactylus planirostris*). *Naturwissenschaften* 97: 781-796.
- Rödder, D., S. Schmidlein, M. Veith, and S. Lötters (2009). Alien invasive slider turtle in unpredicted habitat: a matter of niche shift or of predictors studied? *PloS one* 4, e7843.
- Santika, T., and Hutchinson, M.F. (2009). The effect of species response form on species distribution model prediction and inference. *Ecological Modelling* 220: 2365-2379.
- Saupe, E., V. Barve, C. Myers, J. Soberón, N. Barve, C. Hensz, A.T. Peterson, H.L. Owens, and A. Lira-Noriega (2012). Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling* 237–238: 11-22.
- Soberón, J., and A.T. Peterson (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* 2: 1–10.
- Swets, K. (1988). Measuring the accuracy of diagnostic systems. *Science* 240: 1285-1293.
- Synes, N.W., and P.E. Osborne (2011). Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Global Ecology and Biogeography* 20: 904-914.
- Thomas, C.D., A. Cameron, R.E. Green, M. Bakkenes, L.J. Beaumont, Y.C. Collingham, B. F.N. Erasmus, M. Ferreira de Siqueira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A.S. van Jaarsveld, G.F. Midgley, L. Miles, M.A. Ortega-Huerta, A.T. Peterson, O.L. Phillips, and S.E. Williams (2004). Extinction risk from climate change. *Nature* 427: 145-148.
- Thuiller, W., B. Lafourcade, R. Engler, and M.B. Araújo (2009). BIOMOD - A platform for ensemble forecasting of species distributions. *Ecography* 32: 369-373.
- Varela, S., J.M. Lobo, and J. Hortal (2011). Using species distribution models in palaeobiogeography: A matter of data, predictors and concepts. *Palaeogeography, Palaeoclimatology, Palaeoecology* 310: 451-463.
- Wisn, M.S., R.J. Hijmans, J. Li, A.T. Peterson, C.H. Graham, A. Guisan, and NCEAS Predicting Species Distributions Working Group (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions* 14: 763-773.