



A single-algorithm ensemble approach to estimating suitability and uncertainty: cross-time projections for four Malagasy tenrecs

Robert A. Boria^{1*}, Link E. Olson², Steven M. Goodman^{3,4} and Robert P. Anderson^{1,5,6}

¹Department of Biology, City College of New York, City University of New York, New York, NY 10031, USA, ²University of Alaska Museum, Fairbanks, AK 99708, USA, ³Field Museum of Natural History, Chicago, IL 60605, USA, ⁴Association Vahatra, BP 3972, Antananarivo 101, Madagascar, ⁵Graduate Center, City University of New York, New York, NY 10016, USA, ⁶Division of Vertebrate Zoology (Mammalogy), American Museum of Natural History, New York, NY 10024, USA

ABSTRACT

Aim Ecological niche models (ENMs) are used widely in ecology, evolution, global change biology, but model uncertainty remains an underappreciated issue. Generally, either a single model from one algorithm or an ensemble of single models from different algorithms is used to provide a prediction. In addition to variability among algorithms, recent studies have shown the need to consider variability within a single algorithm, for example optimizing model complexity by tuning model settings. We present an ensemble ENM using a single-algorithm approach, while adjusting model settings to maximize performance.

Location Madagascar.

Methods We used MAXENT, bioclimatic variables and occurrence records of four species of Malagasy tenrecs (Family Tenrecidae). We calibrated and evaluated preliminary models using a jackknife approach, tuning two model settings to estimate optimal model complexity. We chose a suite of top-performing preliminary models and then generated a consensus prediction. Furthermore, we calculated the variability among predictions of the co-optimal models to indicate variation in geography (i.e. uncertainty). We then did the same after projecting the predictions to climatic estimates for the Last Glacial Maximum and the year 2070.

Results The default settings were never identified as optimal for any of the four species. The model settings considered as the co-optimal solutions essentially led to the same evaluation statistics; however, they showed high variation in their geographic predictions for three of the four species. Additionally, variation among such models was greater when transferred across time.

Main conclusions This approach likely can provide better predictions for a single algorithm as well as quantifications of within-algorithm uncertainty, qualities that are highly useful in interpreting reconstructed suitable areas or forecasts of potential range shifts under future climate change. Finally, this within-algorithm uncertainty can be integrated into a larger framework that considers variability due to other factors (e.g. related to input data, alternate algorithms or various Global circulation models).

Keywords

climate change, ecological niche model, ensemble, forecasting, hindcasting, MAXENT, *Microgale*, model complexity, species distribution model, uncertainty.

*Correspondence (current address): Robert A. Boria, University of California, Merced; 5200 North Lake Road, Science and Engineering Building 1, Merced, CA 95343, USA.
E-mail: rboria@ucmerced.edu

INTRODUCTION

Correlative ecological niche models (ENMs), often also termed species distribution models (SDMs), aim to approximate the abiotically suitable areas for a species. They do so by comparing environmental conditions at localities where the species occurs with the overall range of conditions available in the study region (or those unoccupied by the species; Peterson *et al.*, 2011; Anderson, 2012). Such models hold important uses when transferred across time and space, applications that require model generality. The most frequent implementations of ENMs are algorithms that use occurrence localities and either pseudo-absences or background localities (Elith *et al.*, 2006).

Although correlative niche models are used widely in ecology, evolution and global change biology, several outstanding issues remain (Anderson, 2012, 2013), including model uncertainty. There can be many different sources of uncertainty, including algorithm choice and parameterization. Quantification of uncertainty is crucial given that its effects extend to downstream applications of the models (e.g. future climatic estimates based on global circulation models and different emissions scenarios). Whereas uncertainty due to algorithm choice has been explored in previous studies (e.g. Diniz-Filho *et al.*, 2009; Garcia *et al.*, 2012; Watling *et al.*, 2015), that pertaining to model parameterization has seldom been addressed, despite the call for such studies (Araújo & Guisan, 2006).

Here, we explore the effects of algorithm parameterization (i.e. model tuning or smoothing) and its associated maps of uncertainty, with a special emphasis on inference across different periods (past and future). A primary challenge in niche modelling is to balance generality with complexity, especially if the desired outcome is to extrapolate across time or space (Peterson *et al.*, 2011; Anderson, 2013). Recently, several studies have demonstrated the need to tune model settings for a single algorithm (e.g. for the common presence-background technique `MAXENT`, Anderson & Gonzalez, 2011; Elith *et al.*, 2011; Merow *et al.*, 2013). However, most studies still rely on default settings during the model-building stage (reviewed by Merow *et al.*, 2013), even though those settings have been shown to produce overfit models in several instances (Warren & Seifert, 2011; Shcheglovitova & Anderson, 2013; Radosavljevic & Anderson, 2014; Moreno-Amat *et al.*, 2015). Additionally, most studies generally develop a single final model, either by default settings or by optimizing settings, and interpret only that prediction. Measures of algorithmic uncertainty are generally not taken into account when using presence-only models despite the clear need to do so.

An additional approach is to generate an ensemble model, where several model outputs are combined to generate a single prediction. These model outputs can be a combination of the results of different algorithms, different input data (occurrence localities and environmental

variables) and/or different model parameters (Araújo & New, 2007). Frequently, such ensemble modelling has been used for forecasting suitable areas under various climate change scenarios, because it has been shown that different algorithms and future climate estimates can give notably divergent predictions (e.g. Araújo *et al.*, 2005; Pearson *et al.*, 2006; Garcia *et al.*, 2012). Because several different model outputs are combined, measures of uncertainty can be assessed (e.g. Marmion *et al.*, 2009; Garcia *et al.*, 2012). However, despite the fact that varying model parameters has already been proposed for the ensemble approach (Araújo & New, 2007; Alvarado-Serrano & Knowles, 2014), a main drawback when using several different algorithms is that typically the default settings are used for each algorithm. This situation likely derives from impediments for tuning models in current software implementations (Diniz-Filho *et al.*, 2009; Thuiller *et al.*, 2014). Promisingly, some studies have created consensus predictions using models calibrated with different parameters by employing either a Bayesian or multimodel inference approach when using regression-based techniques (Wintle *et al.*, 2003; Hartley *et al.*, 2006). However, these studies focused mainly on variability among models built with different parameters and did not assess variability after projecting the predictions to different periods.

Here, we address these key issues of variability for a prevalent presence-only technique by tuning model settings for a single algorithm and combining the best-performing models using an ensemble approach to generate a consensus prediction and maps of uncertainty. Our goal was to identify highly performing model settings and determine the degree to which resulting predictions differed in geographic space (during current climatic conditions and projections across time). This approach would allow modelling the species' potential distribution and providing a reasonable realm of variability for a given species. As an example, we study four closely related species of Malagasy tenrecs and used the presence-only algorithm `MAXENT` (Phillips *et al.*, 2006; Phillips & Dudík, 2008). In addition to assessing variability in the present, we also examine the ramifications for climate change studies by projecting the three models to climatic estimates for the Last Glacial Maximum (21 kBP) and the year 2070. We know of no prior study that has evaluated the differences among co-optimal solutions for `MAXENT`, or of any that has used a single algorithm to generate ensemble predictions and assess these differences across different periods (but see Moreno-Amat *et al.*, 2015 for differences in model performance when projecting models through time). We predict that the co-optimal solutions will vary in geography, but the extent of that variation is unknown and likely differs among species. Furthermore, theory suggests that such differences in geography will be greater when the models are transferred to another place or period (Peterson *et al.*, 2011).

METHODS

Input data

We conducted analyses for four species found in generally similar habitats, but that vary in body size, locomotion and available sample sizes. Madagascar is home to four endemic radiations of extant terrestrial mammals: lemurs, euplerid carnivorans, nesomyine rodents and tenrecs. The latter have evolved multiple ecomorphologies, but the most species-rich group is the shrew tenrecs of the genus *Microgale*. This genus holds 23 currently recognized extant species (Goodman *et al.*, 2006; Olson *et al.*, 2009; Soarimalala & Goodman, 2011; Olson, 2013), many of which are broadly distributed in the remaining areas of the island's forests. Here, we modelled the abiotically suitable areas for four shrew tenrecs living in humid forest habitats using variable numbers of occurrence records: *Microgale cowani* ($n = 55$ localities), *Microgale gracilis* ($n = 20$), *Microgale gymnorhyncha* ($n = 31$) and *Microgale thomasi* ($n = 28$). These four species are terrestrial faunivores and vary in average body mass from about 11 to 22 g. While *M. cowani* and *M. thomasi* appear to be locomotory generalists, *M. gracilis* and *M. gymnorhyncha* are thought to be at least partially fossorial (Soarimalala & Goodman, 2011). The occurrence localities and their associated georeferences for each species were compiled from field collections and associated notes and the literature (Appendix S1 in Supporting Information). We extended the modelling results of a previously published study for all tenrecs (Goodman *et al.* (2013), which used somewhat different locality datasets) using more refined modelling methods. The variety of morphological and natural history traits of these species likely lead to differences in demography and dispersal ability, allowing us to consider variation among species that generally inhabit similar environmental conditions.

Generally, researchers tend to sample more easily accessible areas, leading to localities that are geographically clustered (Hijmans *et al.*, 2000; Reddy & Dávalos, 2003; Kadmon *et al.*, 2004). This can artificially increase spatial autocorrelation of the resulting localities for a given species and bias the niche signal. Eliminating artefactual clusters of localities is also important for model evaluation, as evaluation localities that lie close to calibration ones lead to inflated values of performance when sampling is geographically biased (Veloz, 2009; Hijmans, 2012). To reduce the effects of sampling bias, we spatially filtered the occurrence dataset to ensure that no two localities were within 10 km of one another (while retaining the most localities possible) using the `SPTHIN` package in R (Aiello-Lammens *et al.*, 2015, R Development Core Team, 2014; we used a beta version of the package, but the results were not affected). This distance was chosen because a recent study showed that spatially filtering a *M. cowani* dataset by a 10-km distance reduced overfitting and improved model performance (Boria *et al.*, 2014). Because all four species share similar geographic distributions and

result from the same sampling efforts for the target group of small non-volant mammals, we used the 10-km distance for each species. Nevertheless, reducing the datasets through spatial filtering can have at least two unfortunate consequences. First, smaller sample sizes per se can decrease the level of model complexity that is reasonable (and selected via tuning analyses). Secondly, the eliminated localities might document the species' presence in environmental conditions not included in any of the retained records. The latter possibility is particularly likely in areas of highly dissected terrain, such as mountainous portions of Madagascar. Nevertheless, following the large literature regarding sampling bias and its adverse effects on niche models, we conducted spatial filtering with the intent of gaining a more representative, albeit smaller, dataset for each species (Anderson, 2012). Each dataset was reduced after applying the spatial filter: *M. cowani* ($n = 32$), *M. gracilis* ($n = 15$), *M. gymnorhyncha* ($n = 20$) and *M. thomasi* ($n = 17$).

The present-day environmental data were obtained from the WorldClim dataset (<http://www.worldclim.org>; Hijmans *et al.*, 2005; at 30" resolution). These 19 bioclimatic variables reflect aspects of temperature and precipitation and have been used successfully for niche models of many small mammals, including *M. cowani* (Boria *et al.*, 2014). Although these variables can be correlated, `MAXENT` is a machine-learning algorithm that employs regularization to reduce complexity (especially regarding correlated variables), and not all variables are necessarily included in the final model (for details; Phillips & Dudík, 2008; Elith *et al.*, 2011). To approximate modelling assumptions regarding biotic interactions and dispersal/demography more closely, we delimited a custom study region for each species, specifically by drawing a minimum convex polygon around the localities and adding a 0.5° buffer (Anderson & Raza, 2010; Barve *et al.*, 2011; Anderson, 2013). Background localities for calibration were taken from only within the delimited study region.

The reconstructed past climatic conditions (Last Glacial Maximum; 21 kBP; 2.5' resolution) and the projected future ones (2070; 30" resolution; emission scenario rcp2.6) were generated by the community climate system model (CCSM; Kiehl & Gent, 2004; Gent *et al.*, 2011). These layers consisted of the same 19 environmental variables as the current climatic conditions and were downloaded from WorldClim (<http://worldclim.org/>). For an example of transfer across time, we only used these two climate scenarios, but note that great variability also can exist among climate modelling groups and (for future projections) among emissions scenarios (Garcia *et al.*, 2012).

Ecological niche modelling

We calibrated models using the presence-background technique in `MAXENT` (Phillips *et al.*, 2006; Phillips & Dudík, 2008), which has performed well in comparison with other algorithms (Elith *et al.*, 2006; Wisz *et al.*, 2008; Merow &

Silander, 2014; but see Royle *et al.*, 2012; Fitzpatrick *et al.*, 2013). The default settings (= Auto features) of that software were determined based a broad study of a large number of species (with many different sample sizes) from many different parts of the world (Phillips & Dudík, 2008). However, for several reasons, the default settings are expected (and have been shown) to produce overfit models (Radosavljevic & Anderson, 2014).

Users are able to adjust model settings easily in MAXENT for two critical factors: feature classes (FC) and regularization multiplier (RM). Feature classes determine the flexibility of the allowed modelled responses to the environmental predictor variables (e.g. linear, quadratic, product, ... j). The RM (a single coefficient applied to each of the β_j ; regularization values specific to each feature class j) penalizes models for including additional parameters. Hence, it leads the algorithm to avoid parameters that do not introduce substantial new information, reducing the production of overly complex models (Phillips & Dudík, 2008; Merow *et al.*, 2013; Shcheglovitova & Anderson, 2013; Radosavljevic & Anderson, 2014). When using default settings, sample size determines which FCs are allowed and the level of regularization (Phillips & Dudík, 2008).

To select species-specific model settings approximating optimal levels of complexity, we tuned model settings by varying different combinations of FC and RM and applying a jackknife approach of occurrence localities for evaluating model performance (Shcheglovitova & Anderson, 2013). We used the R package ENMEVAL (Muscarella *et al.*, 2014; we used a beta version of the package, but the results were not affected), and all tuning analyses were conducted with the raw output format of MAXENT (although cumulative or logistic rescalings would not affect the omission rates and area under the curve (AUC) values calculated here). The jackknife procedure is special case of k -fold cross-validation, where $k = n$. Here, a model is calibrated using $n - 1$ localities and evaluated on the withheld locality; this is carried out until all localities have been used as an evaluation record exactly once (Pearson *et al.*, 2007). This technique has been shown to identify better performing and more realistic models built with small sample sizes (Shcheglovitova & Anderson, 2013). For consistency of methods and limitations of the jackknife procedure for large sample sizes (which increases the likelihood that calibration localities and evaluation localities will be close to each other in space), we randomly rarefied the dataset of *M. cowani* from 32 localities to 25 (this differed from the spatial filter in not requiring any spatial criteria). Because of the small number of localities for each species, we limited the FCs allowed to the simplest combinations: Linear; Hinge; Linear and Quadratic; and Linear, Quadratic and Hinge. Furthermore, we employed a range of RM values (0.5–6.0 with 0.5 intervals). This led to a total of 48 unique combinations of FCs and RMs used to model the suitability for each species.

Model tuning and optimality criteria

To identify the optimal and co-optimal combinations of model settings, we evaluated model performance using sequential criteria (Shcheglovitova & Anderson, 2013; Muscarella *et al.*, 2014). To avoid models that are overfit to the calibration data, we first employed a threshold-dependent measure, specifically an omission rate (the proportion of evaluation localities that fell outside of the prediction once converted to a binary one). The omission rate used here was based on the lowest presence threshold (LPT = 0% calibration omission rate; Pearson *et al.*, 2007; = minimum training presence threshold of MAXENT). The LPT rule sets the threshold at the lowest value of the prediction for any pixel that contains a calibration locality (Pearson *et al.*, 2007; Anderson & Gonzalez, 2011) and has an expected omission rate of zero for evaluation localities. The higher the omission rate (above that expected value), the more overfit a model.

Table 1 Summary statistics for the five co-optimal solutions in model-tuning experiments for ecological niche models of four Malagasy species of the genus *Microgale* (*M. cowani*, *M. gracilis*, *M. gymnorhyncha* and *M. thomasi*). Preliminary models were calibrated and evaluated using a jackknifing approach, varying two model settings (feature classes and regularization multiplier). Optimal model complexity was then estimated via a sequential procedure (minimizing overfitting and then maximizing discriminatory ability). For each combination of feature classes and regularization multiplier, results were averaged over jackknife iterations for one measure of overfitting (average omission rates at the lowest presence threshold) and one measure of overall discriminatory ability; $AUC_{\text{evaluation}}$). Note the overall similarity among the co-optimal model settings for each species. See supplementary appendix for comparisons with the 48 settings combinations that showed poorer performance for each species.

Species	Model settings	Omission rate	$AUC_{\text{evaluation}}$
<i>M. cowani</i>	LQ_2.5	0.040	0.883
	LQ_3	0.040	0.882
	LQH_3	0.040	0.881
	LQH_3.5	0.040	0.878
<i>M. gracilis</i>	LQ_3.5	0.040	0.878
	LQ_0.5	0.067	0.920
	LQ_1	0.067	0.918
	LQ_1.5	0.067	0.909
	H_2.5	0.067	0.909
<i>M. gymnorhyncha</i>	LQH_2.5	0.067	0.908
	LQ_0.5	0.050	0.925
	LQH_1	0.050	0.918
	LQH_1.5	0.050	0.917
	H_1.5	0.050	0.916
<i>M. thomasi</i>	H_2	0.050	0.915
	LQ_0.5	0.063	0.888
	LQH_3	0.063	0.885
	LQH_3.5	0.063	0.884
	LQH_2	0.063	0.883
	LQH_1	0.063	0.883

This thresholding rule has been shown to be effective, and perhaps even conservative, for small sample sizes (Shcheglovitova & Anderson, 2013). For each of the different combinations of settings (FC/RM), we obtained an omission rate and then averaged these values across the jackknife iterations (Shcheglovitova & Anderson, 2013).

Because several setting combinations can result in the lowest omission rate, we applied another measure of evaluation, one that maximizes discriminatory ability. Specifically, we ordered the models that showed the lowest omission rate by a secondary (sequential) criterion: the rank-based threshold-independent AUC of the receiver operating characteristic plot. As above for omission rates, we obtained an AUC_{evaluation} score for each iteration and then averaged those values across the jackknife iterations (per combination of FC and RM; Shcheglovitova & Anderson, 2013). We determined the optimal combination for each species (of the 48 assessed here) as the model with the lowest average omission rate and, subsequently, the highest average AUC_{evaluation}. We also used these statistics to determine co-optimal settings that performed almost as well as the optimal ones. Several different solutions can have the same omission rate and very similar AUC_{evaluation} scores. We chose the top-performing 10% of models based on the sequential criteria (see above; of the 48 FC and RM combinations) and defined those five as the co-optimal solutions.

Final models and projections

For each species, we made final models for projection using all filtered localities for the identified optimal and co-optimal settings, as well as for the default ones. The final models were then projected to all of Madagascar (and nearshore islands) for current climatic conditions, reconstructed LGM conditions and predicted conditions in the year 2070. For visualization purposes and calculations of Schoener's *D* statistic, we used the logistic output format [for details about logistic format, see Phillips & Dudík (2008) and Merow *et al.* (2013)]. For comparison purposes, we built a model using all filtered localities and the default settings (FC = Linear, Quadratic and Hinge; and RM = 1.0). We visually inspected these predictions in comparison with those resulting from the ensemble model and we quantitatively evaluated their similarities using Schoener's *D*.

Ensemble models and their associated uncertainty

We then created ensemble models for each species and measured uncertainty in two ways. First, we averaged the five-raster grids from the co-optimal settings, separately for each time period (using the RASTER package in R; Hijmans *et al.*, 2011). For experimental simplicity and because the co-optimal solutions are all high-performing models, we chose to average the rasters (but note that other methods exist for creating consensus models via ensemble modelling, e.g. via unequal weighting; Araújo *et al.*, 2005; Marmion *et al.*, 2009). To obtain a map

Table 2 Pairwise Schoener's *D* statistic measuring the geographic similarity for the five co-optimal solutions in ecological niche models for four Malagasy species of Tenrecidae – *Microgale cowani*, *M. gracilis*, *M. gymnorhyncha* and *M. thomasi*. Note the high level of variability for pairs of co-optimal models made with different settings for most species, but much less so for *M. cowani*.

Model settings	LQ_2.5	LQ_3	LQH_3	LQH_3.5	LQ_3.5
<i>M. cowani</i>					
LQ_2.5	1	–	–	–	–
LQ_3	0.965	1	–	–	–
LQH_3	0.974	0.981	1	–	–
LQH_3.5	0.947	0.980	0.968	1	–
LQ_3.5	0.944	0.977	0.963	0.994	1
	LQ_0.5	LQ_1.0	LQ_1.5	H_2.5	LQH_2.5
<i>M. gracilis</i>					
LQ_0.5	1	–	–	–	–
LQ_1	0.863	1	–	–	–
LQ_1.5	0.802	0.935	1	–	–
H_2.5	0.647	0.748	0.789	1	–
LQH_2.5	0.734	0.846	0.882	0.873	1
	LQ_0.5	LQH_1	LQH_1.5	H_1.5	H_2
<i>M. gymnorhyncha</i>					
LQ_0.5	1	–	–	–	–
LQH_1	0.872	1	–	–	–
LQH_1.5	0.863	0.930	1	–	–
H_1.5	0.812	0.872	0.901	1	–
H_2	0.787	0.824	0.867	0.936	1
	LQ_0.5	LQH_3	LQH_3.5	LQH_2	LQH_1
<i>M. thomasi</i>					
LQ_0.5	1	–	–	–	–
LQH_3	0.744	1	–	–	–
LQH_3.5	0.692	0.936	1	–	–
LQH_2	0.820	0.898	0.837	1	–
LQH_1	0.863	0.763	0.711	0.849	1

showing variation among co-optimal predictions in geography, we calculated the standard deviation of the predictions using the RASTER package in R (= uncertainty map; Hijmans *et al.*, 2011). Second, as another approach to measuring uncertainty and because many studies employ thresholding rules, we converted each of the co-optimal models from the current climatic conditions into binary predictions. Specifically, we again used the LPT threshold and then summed the binary maps for each species. The values of these maps range from 0 (where all of the models predict no suitable areas) to 5 (where all of the predictions indicate suitable areas).

To determine variability in geography among models, we measured the pairwise similarity by calculating Schoener's *D* values using the R package ENMEVAL (Warren *et al.*, 2009; Muscarella *et al.*, 2014). Schoener's *D* is a pixel-by-pixel comparison between two predictions with scores ranging

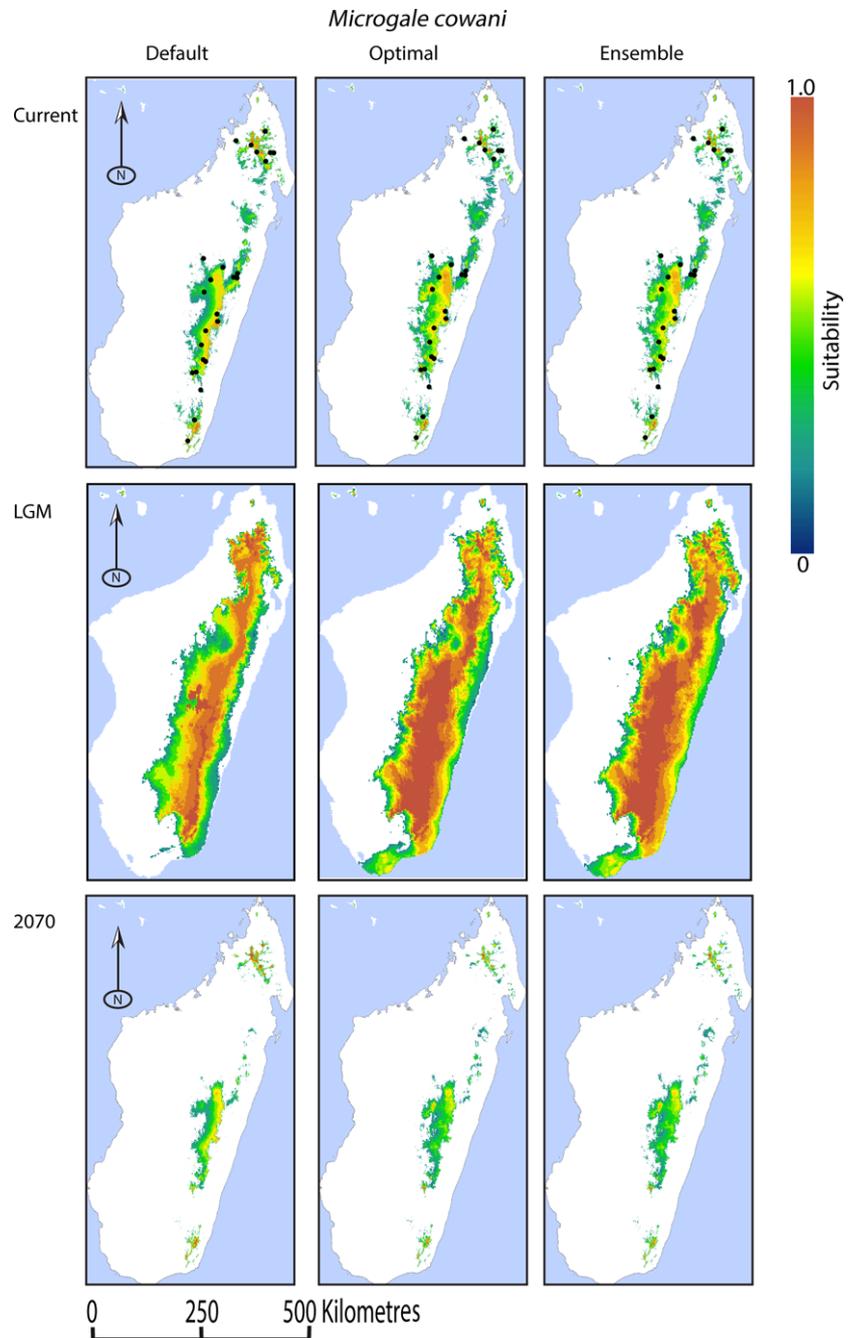


Figure 1 Maps showing three different ecological niche models for *Microgale cowani*: those based on default settings (left), optimal settings (middle) and an ensemble of co-optimal settings (right). Results are shown for three different time slices: current (top), the Last Glacial Maximum (=LGM; cooler and drier conditions; 21 kBP; middle) and the year 2070 (bottom; warmer and drier). Increasingly warm colours indicate successively stronger predictions (higher suitability). Areas in white indicate pixels with values below the LTP threshold determined based on calibration localities. Note that despite the overall high similarity between predictions based on the optimal settings and the ensemble of co-optimal settings, the ensemble prediction shows a broader extent for each time slice. Colour figure can be viewed at wileyonlinelibrary.com

from 0 (no overlap) to 1 (identical models) (Warren *et al.*, 2009). We did so among the co-optimal solutions when projected to each time period. Additionally, to assess similarity among the three different single models (default, optimal and ensemble) for the current climatic conditions, we calculated pairwise Schoener's D_s between them.

RESULTS

Model tuning

Using the sequential criteria, the model-tuning exercises yielded several clear trends regarding RM and the evaluation

statistics. For all FCs, omission rate generally decreased with increasing RM for *M. cowani* and *M. gracilis* but lacked consistent trends for *M. gymnorhyncha* and *M. thomasi* (Fig. S1a,c,e,g). Additionally, for each FC, $AUC_{\text{evaluation}}$ almost always showed a steady decreasing trend with increasing RMs (Fig. S1b,d,f,h).

The default settings were never identified as optimal for any of the four species. The optimal model settings (FC + RM) were the same for three species (*M. gracilis*, *M. gymnorhyncha* and *M. thomasi*): Linear and Quadratic with a RM of 0.5; furthermore, the optimal setting for *M. cowani* contained the same FC combination with a RM of 2.5 (Table 1). The Linear and Quadratic FCs allow less

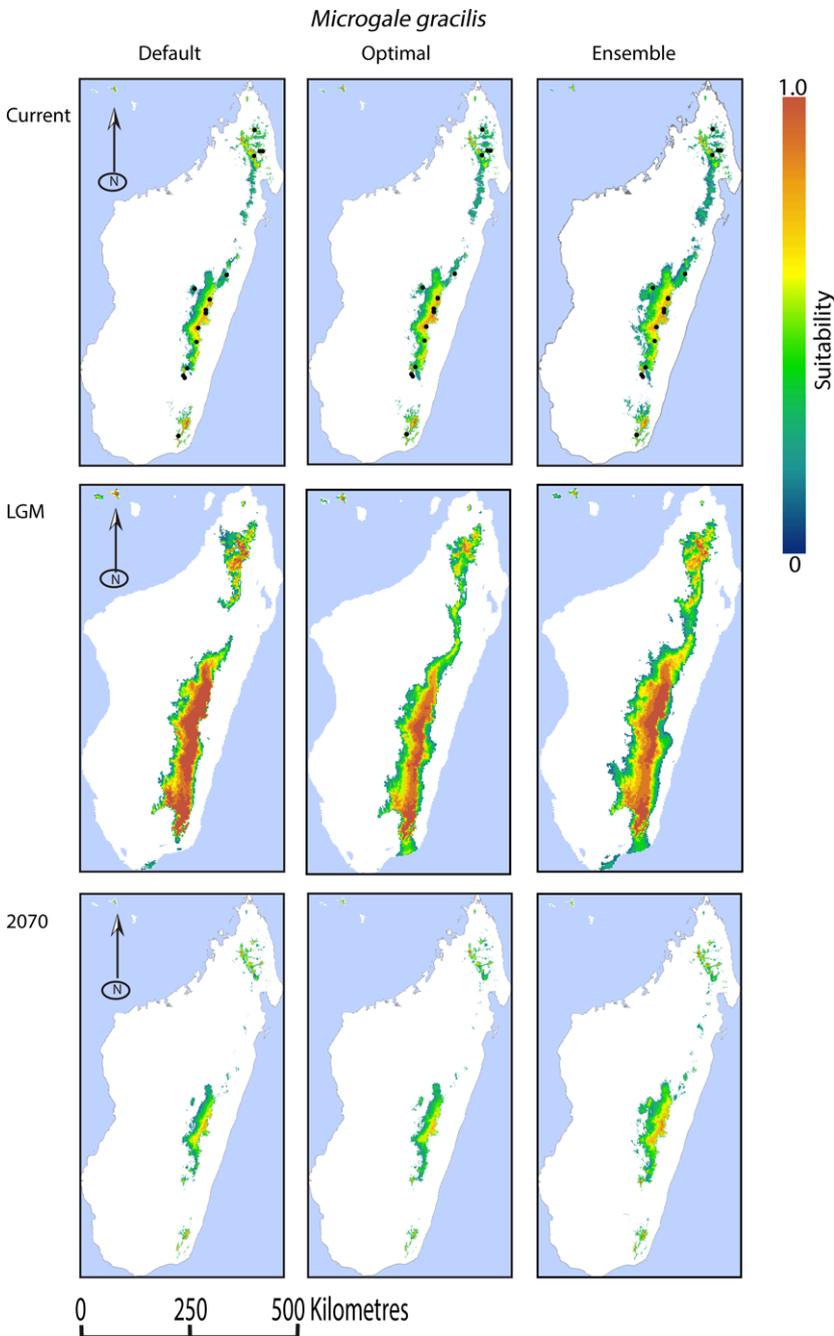


Figure 2 Maps showing the three different ecological niche models for *Microgale gracilis*: those based on default settings (left), optimal settings (middle) and an ensemble of co-optimal settings (right). Results are shown for three different time slices: current (top), the Last Glacial Maximum (=LGM; 21 kBP; middle) and the year 2070 (bottom). Increasingly warm colours indicate successively stronger predictions (higher suitability). Areas in white indicate pixels with values below the LTP threshold determined based on calibration localities. Note the overall strong differences between the ensemble model and the other two prediction approaches, especially for the LGM predictions. Colour figure can be viewed at wileyonlinelibrary.com

complex responses than the default settings, which also included the Hinge FC. However, in a compensatory fashion, the 0.5 RM was smaller than the default setting. In contrast, the 2.5 RM is higher than the default setting (penalizing complexity; Table 1).

The five model settings considered as the co-optimal solutions essentially led to the same evaluation statistics (Table 1). For each species, several model-setting combinations achieved the lowest average omission rate, and many of those combinations had $AUC_{\text{evaluation}}$ scores that were very similar to the respective optimal model (Fig. S1a–h; Table 1). Many of the co-optimal solutions consisted of Linear and Quadratic or Linear, Quadratic and Hinge FCs. Only two

species included co-optimal model combinations that used only the Hinge feature class: *M. gracilis* (one of the co-optimal combinations) and *M. gymnorhyncha* (two of the co-optimal combinations). The co-optimal RMs ranged from 0.5 to 3.5. Finally, the default settings were not part of the co-optimal solutions for *M. cowani* and *M. gracilis*, but they were identified as one of the five co-optimal settings for *M. gymnorhyncha* and *M. thomasi*.

Ensemble models and maps of uncertainty

The co-optimal solutions differed substantially in geography for three of the four species (Table 2). For *Microgale gracilis*,

Table 3 Pairwise Schoener's *D* statistic measuring the geographic similarity of three different ecological niche models: default settings, optimal settings and ensemble prediction for each of four Malagasy shrew tenrecs. Note the high level of variability among the three different models for all four species.

Model	Default	Optimal	Ensemble
<i>M. cowani</i>			
Default	1	–	–
Optimal	0.804	1	–
Ensemble	0.805	0.976	1
<i>M. gracilis</i>			
Default	1	–	–
Optimal	0.886	1	–
Ensemble	0.820	0.780	1
<i>M. gymnorhyncha</i>			
Default	1	–	–
Optimal	0.873	1	–
Ensemble	0.909	0.862	1
<i>M. thomasi</i>			
Default	1	–	–
Optimal	0.863	1	–
Ensemble	0.823	0.803	1

M. gymnorhyncha and *M. thomasi*, Schoener's *D* consistently showed a larger range in variation. In contrast, the co-optimal solutions for *M. cowani* yielded very similar predictions in geography. For illustration of each category, we present geographic predictions for one species that varied greatly (*M. gracilis*) and the one species that showed little variation (*M. cowani*; Figs 1 & 2).

The predictions for current climatic conditions for *M. cowani* and *M. gracilis* for the three different models (default, optimal and ensemble) showed clear quantitative differences in geographic space (Table 3). The default settings for *M. cowani* led to predictions that were more different than those for the optimal and ensemble models. In contrast, the default settings prediction for *M. gracilis* was more similar to that for the optimal settings than to the one for the ensemble model. Qualitative evaluations among the predictions in geographic space for the three time periods revealed several key differences, especially for *M. gracilis* (Fig. 1; and see also *M. gymnorhyncha* and *M. thomasi*; Figs S2 & S3). Inspection of the three types of predictions indicated differences that ranged from minimal (*M. cowani*) to very strong (*M. gracilis*, *M. gymnorhyncha* and *M. thomasi*) in the extent of predicted areas for each of the three periods (Figs 1 & 2; Figs S2 & S3). For each period, the default settings for *M. cowani* led to a more restricted prediction than did either the optimal or ensemble predictions (Fig. 1). Interestingly, for *M. gracilis*, the optimal solution was more similar to the default settings for two of the three time slices than the ensemble prediction. Again, for that species, the ensemble model predicted a larger geographic extent of suitable areas than the other two models did. The most striking difference for *M. gracilis* appeared for the LGM predictions. The default settings predicted suitable areas as

disjunct, while the optimal settings and the ensemble models predicted a contiguous suitable area throughout Madagascar's eastern humid forest (Fig. 2; see also Figs S2 & S3). For all four species, the predictions for every model indicated smaller potential suitable areas for the year 2070 (last row of Figs 1 & 2; Figs S2 & S3).

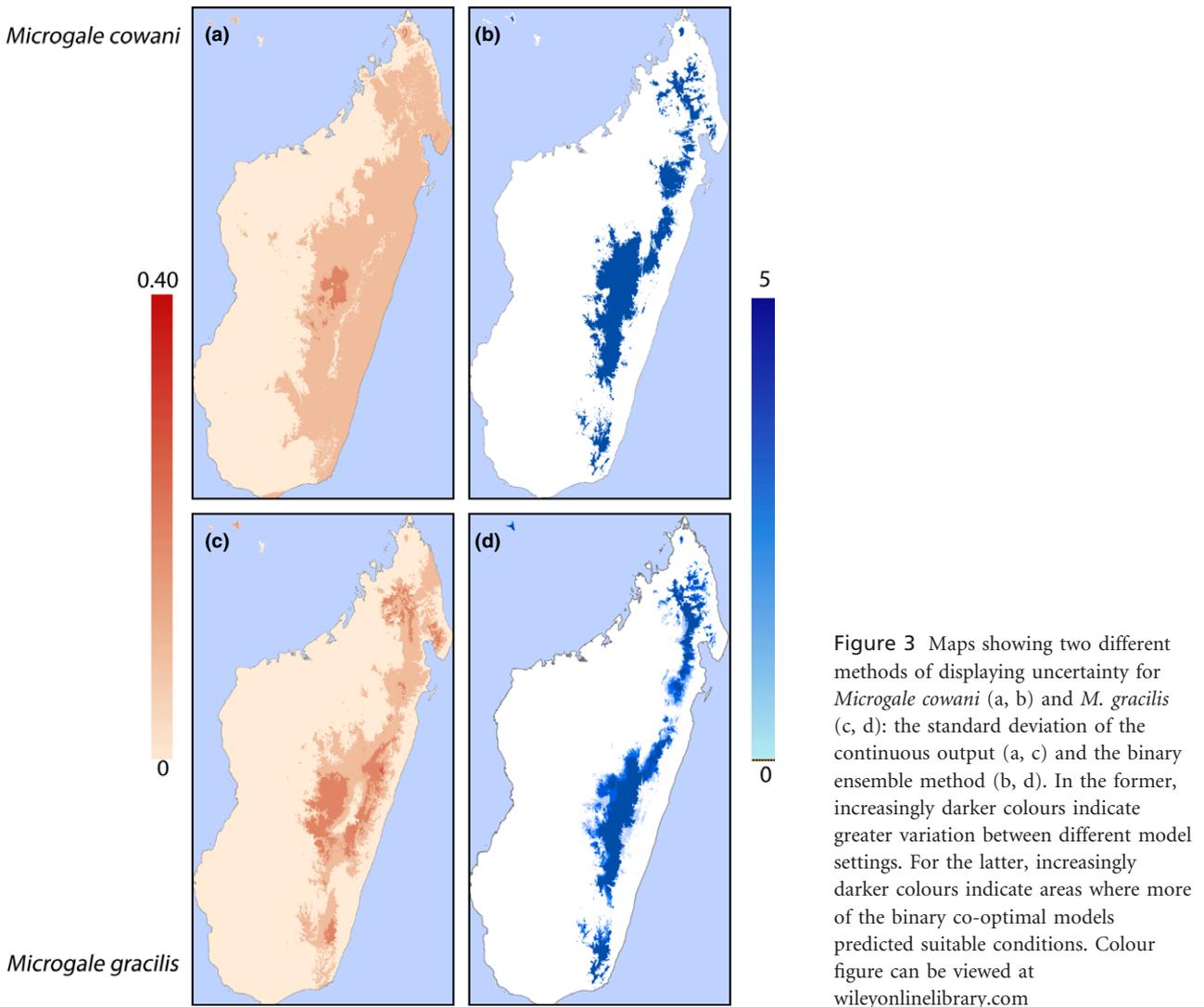
The maps of uncertainty show interesting patterns of variation among co-optimal solutions in geographic space for the three time slices (Fig. 3; Figs S4 & S5). First, the standard deviation maps revealed that the areas of the predictions were highly variable among the model settings differed among species. For example, *M. cowani* showed the most variation in central Madagascar, and *M. gracilis* exhibited considerable variation in the mid-latitude and northern portions of the island (Fig. 3). For both species, there was little variation in areas that are predicted as unsuitable, specifically the arid western portions of Madagascar, which host far fewer species of shrew tenrecs than the eastern humid forests (Soarimalala & Goodman, 2011; Goodman *et al.*, 2013). Further, there was little variation for *M. cowani* when projected to different time periods, but *M. gracilis* displayed greater discrepancies (Fig. S5, see also the tremendous variation for *M. gymnorhyncha* and *M. thomasi* at LGM). Additionally, the summed binary models for each species indicated that once converted into a binary prediction, many of the co-optimal solutions (for current climatic conditions) predict very similar areas as suitable (Fig. 3; Fig. S4). The specific areas of high variability according to the standard deviation map were also areas of discordance for the binary predictions for *M. gracilis* (and *M. thomasi*).

DISCUSSION

Importance of model tuning and co-optimal solutions

The default settings were not selected as optimal for any of the four species (consistent with results from other recent studies with MAXENT; Warren & Seifert, 2011; Shcheglovitova & Anderson, 2013; Radosavljevic & Anderson, 2014), but in two cases the default settings were among the co-optimal ones. The co-optimal solutions had essentially the same evaluation statistics (Table 1), with less than a 0.01 difference in their AUC_{evaluation} scores. Although the default settings were not part of the co-optimal solutions for *M. cowani* and *M. gracilis*, they were identified as a co-optimal setting for *M. gymnorhyncha* and *M. thomasi*. Hence, for some species, the default settings may lead to models that are not overfit, as originally intended by Phillips & Dudík (2008).

Although the co-optimal solutions were very similar in their evaluation statistics, the settings for three of the four species led to models that varied considerably in geographic space (Table 2; Fig. 3). For example, the co-optimal solutions for *M. gracilis* were only 81% similar on average according to Schoener's *D* (Table 2). Additionally, among the five co-optimal solutions for *M. gracilis*, only one used



Hinge FC. All of the other co-optimal settings were either Linear and Quadratic, or Linear, Quadratic and Hinge. For this species, there were several different mathematical solutions to construct quantitatively similarly performing models that nevertheless differed in geographic space. When the thresholding rule was applied to models projected to the LGM for *M. gracilis*, we identified key differences among the different model types (Fig. 2). Finally, the maps of uncertainty indicated the areas in geographic space that vary markedly with regard to suitability (Fig. 3; Figs S4 & S5). This is especially clear when no threshold rule is applied to the predictions (Table 2).

Future climate change projections

All four Malagasy tenrecs would lose abiotically suitable areas under the future projections employed here (bottom row of Figs 1 & 2; Figs S2 & S3). Madagascar has already lost about 90% of its estimated natural forest due to deforestation (Harper *et al.*, 2007). Currently, all four species are

classified by the IUCN Red List as Least Concern (Stephenson *et al.*, 2016a,b,c,e), but in the case of *M. cowani* (and possibly one or more of the other species), ongoing taxonomic revision will likely identify additional constituent 'cryptic' species with more restricted ranges, rendering our predictions conservative (L.E. Olson, unpublished data). This has already been borne out in the case of *M. jobihely*, a 'cryptic species' formerly confused with *M. cowani* and described only a decade ago (Goodman *et al.*, 2006); *M. jobihely* is now classified as endangered by the IUCN (Stephenson *et al.*, 2016d) due to its restricted range. Furthermore, the current analyses only include one scenario for future temperature and precipitation. A more comprehensive study should be conducted that includes more climatic scenarios (e.g. Garcia *et al.*, 2012), uses forest cover (e.g. Feddema *et al.*, 2005) and includes demographic information to simulate dispersal scenarios to predict future occupied areas (e.g. Fordham *et al.*, 2012). A final aspect that needs to be taken into consideration is tolerance to unfor- ested areas. Recent field surveys have found *M. cowani* in

agricultural zones about 500 m from the nearest forest habitat and *M. thomasi* in heavily degraded secondary forest habitat (Randriamoria *et al.*, 2015). Hence, for these two species, evidence exists indicating that they can tolerate some anthropogenic alterations.

Conclusions and recommendations for transfers across time and space

This preliminary study using one algorithm, MAXENT, illustrates the benefits of making an ensemble prediction by tuning model settings to select multiple high-performing combinations. Using this single-algorithm ensemble approach, we found that co-optimal solutions varied in geographic space (Table 2). Although the predictions were visually fairly similar in current climatic conditions, the differences became exacerbated as models were projected through time. The variability within a single algorithm (most notably with different model settings) within studies that utilize ENMs is rarely assessed (see Wintle *et al.*, 2003; Hartley *et al.*, 2006), but the current study suggests that it can substantially affect the results and conclusions.

There are several caveats in the current study that should be addressed in follow-up studies. First, using different geographic filtering distances may allow inclusion of more localities, and potentially add valuable environmental information. Second, the jackknife approach does not test the performance of a model after transfers across space, which should be done when the desired use is to transfer across periods (Araújo & Guisan, 2006). Preferably, researchers will have a sufficient number of occurrence records to perform tests of spatially independent evaluations (Boria *et al.*, 2014; Radosavljevic & Anderson, 2014); unfortunately, however, most species are poorly documented and a lack of data limits these tests. Third, we did not fully explore parameter space with MAXENT; therefore, other studies could determine whether using a greater variety of settings yields greater similarity among predictions. Fourth, studies could employ other criteria and methods to generate the co-optimal solutions, such as using AIC (Warren & Seifert, 2011). Fifth, the number of co-optimal solutions can be varied, and concordance among the settings can be investigated. In addition to refining the current methodology, other factors could also be explored (e.g. variable selection, performance metric and studying differences in environmental rather than geographic space). Despite these caveats, these results demonstrate that different model settings, even if they show similar quantitative performance, can generate vastly different reconstructions of past suitable areas or future projections, variation that should be taken into account in making biological inferences (Fig. 2; Figs S2 & S3).

This approach is also scalable (i.e. including different algorithms and more parameter space). Future studies should be undertaken to test whether these results hold across multiple species in different geographic regions. As mentioned earlier, several different algorithms (also built with different climate

hind/forecasts) can be parameterized independently and their co-optimal outputs combined to generate a super-ensemble of tuned models, as originally envisioned by Araújo & New (2007).

ACKNOWLEDGEMENTS

This research was made possible by funding from the U.S. National Science Foundation (NSF DEB-1119915 to R.P.A. and DEB-1120904 to L.E.O.) and the Luis Stokes Alliance for Minority Participation (Bridge to Doctorate Fellowship to R.A.B.). Jason L. Brown provided environmental layers for the Last Glacial Maximum. We thank Matthew Aiello-Lammens, Diego Alvarado-Serrano, Babatunde Alade, Amy C. Berkov, Jason L. Brown, Ana C. Carnaval, Peter J. Galante, Maria Gavrutenko, Beth E. Gerstner, Jamie M. Kass, Mariano Soley-Guardia and two anonymous referees for their helpful comments and/or discussion on one or more versions of the manuscript.

REFERENCES

- Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B. & Anderson, R.P. (2015) spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, **38**, 541–545.
- Alvarado-Serrano, D. & Knowles, L.L. (2014) Ecological niche models in phylogeographic studies: applications, advances and precautions. *Molecular Ecology Resources*, **14**, 233–248.
- Anderson, R.P. (2012) Harnessing the world's biodiversity data: promise and peril in ecological niche modelling of species distributions. *Annals of the New York Academy of Sciences*, **1260**, 66–80.
- Anderson, R.P. (2013) A framework for using niche models to estimate impacts of climate change on species distributions. *Annals of the New York Academy of Sciences*, **1297**, 8–28.
- Anderson, R.P. & Gonzalez Jr, I. (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with MaxEnt. *Ecological Modelling*, **222**, 2796–2811.
- Anderson, R.P. & Raza, A. (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*, **37**, 1378–1393.
- Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677–1688.
- Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.
- Araújo, M.B., Whittaker, R.J., Ladle, R.J. & Erhard, M. (2005) Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, **14**, 529–538.

- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modelling and species distribution modelling. *Ecological Modelling*, **222**, 1810–1819.
- Boria, R.A., Olson, L.E., Goodman, S.M. & Anderson, R.P. (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, **275**, 73–77.
- Diniz-Filho, J.A.F., Bini, L.M. & Rangel, T.F. (2009) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography*, **32**, 897–906.
- Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.
- Feddema, J.J., Oleson, K.W., Bonan, G.B., Mearns, L.O., Buja, L.E., Meehl, G.A. & Washington, W.M. (2005) The importance of land-cover change in simulating future climates. *Science*, **310**, 1674–1678.
- Fitzpatrick, M.C., Gotelli, G.J. & Ellison, A.M. (2013) MaxEnt versus MaxLike: empirical comparisons with ant species distributions. *Ecosphere*, **4**, 1–15.
- Fordham, D.A., Resit Akçakaya, H., Araújo, M.B., Elith, J., Keith, D.A., Pearson, R., Auld, T.D., Mellin, C., Morgan, J.W., Regan, T.J., Tozer, M., Watts, M.J., White, M., Wintle, B.A., Yates, C. & Brook, B.W. (2012) Plant extinction risk under climate change: are forecast range shifts alone a good indicator of species vulnerability to global warming? *Global Change Biology*, **18**, 1357–1371.
- García, R.A., Burgess, N.D., Cabeza, M., Rahbek, C. & Araújo, M.B. (2012) Exploring consensus in 21st century projections of climatically suitable areas for African vertebrates. *Global Change Biology*, **18**, 1253–1269.
- Gent, P.R., Danabasoglu, G., Donner, L.J., Holland, M.M., Hunke, E.C., Jayne, S.R., Lawrence, D.M., Neale, R.B., Rasch, P.J., Vertenstein, M., Worley, P.H., Yang, Z.-L. & Zhang, M. (2011) The community climate system model version 4. *Journal of Climate*, **24**, 4973–4991.
- Goodman, S.M., Raxworthy, C.J., Maminirina, C.P. & Olson, L.E. (2006) A new species of shrew tenrec (*Microgale jobihely*) from northern Madagascar. *Journal of Zoology*, **270**, 384–398.
- Goodman, S.M., Soarimalala, V., Raheriarisena, M. & Rakotondravony, D. (2013) Small mammals or tenrecs (Tenrecidae) and rodents (Nesomyidae). *Atlas of selected land vertebrates of Madagascar* (ed. by S.M. Goodman and M.J. Raherilalao), pp. 211–269. Association Vahatra, Antananarivo.
- Harper, G.J., Steininger, M.K., Tucker, C.J., Juhn, D. & Hawkins, F. (2007) Fifty years of deforestation and forest fragmentation in Madagascar. *Environmental Conservation*, **34**, 325–333.
- Hartley, S., Harris, R. & Lester, P.J. (2006) Quantifying uncertainty in the potential distribution of an invasive species: climate and the Argentine ant. *Ecology Letters*, **9**, 1068–1079.
- Hijmans, R.J. (2012) Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology*, **93**, 679–688.
- Hijmans, R.J., Garrett, K.A., Huaman, Z., Zhang, D.P., Schreuder, M. & Bonierbale, M. (2000) Assessing the geographic representativeness of genebank collections: the case of the Bolivian wild potatoes. *Conservation Biology*, **14**, 1755–1765.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2011) dismo package for R version 1.1-1. Available at: <http://cran.r-project.org/web/packages/dismo/index.html>.
- Kadmon, R., Farber, O. & Danin, A. (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, **14**, 401–413.
- Kiehl, J.T. & Gent, P.R. (2004) The Community Climate System Model, Version 2 (CCSM2). *Journal of Climate*, **17**, 3666–3682.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, **15**, 59–69.
- Merow, C. & Silander, J.A. (2014) A comparison of Maxlike and Maxent for modelling species distributions. *Methods in Ecology and Evolution*, **5**, 215–225.
- Merow, C., Smith, M.J. & Silander, J.A. (2013) A practical guide to MaxEnt for modelling species' distributions: what it does, and why inputs and settings matter. *Ecography*, **36**, 1058–1069.
- Moreno-Amat, E., Mateo, R.G., Nieto-Lugilde, D., Morueta-Holme, N., Svenning, J.C. & García-Amorena, I. (2015) Impact of model complexity on cross-temporal transferability in Maxent species distribution models: an assessment using paleobotanical data. *Ecological Modelling*, **312**, 308–317.
- Muscarella, R., Galante, P.J., Soley-Guardia, M., Boria, R.A., Kass, J.M., Uriarte, M. & Anderson, R.P. (2014) ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution*, **5**, 1198–1205.
- Olson, L.E. (2013) Tenrecs. *Current Biology*, **23**, R5–R8.
- Olson, L.E., Rakotomalala, Z., Hildebrandt, K.B.P., Lanier, H.C., Raxworthy, C.J. & Goodman, S.M. (2009) Phylogeography of *Microgale breviceaudata* (Tenrecidae) and description of a new species from western Madagascar. *Journal of Mammalogy*, **90**, 1095–1110.

- 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.
- Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Peterson, A.T. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.
- Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martinez-Meyer, E., Nakamura, M. & Araújo, M.B. (2011) *Ecological niches and geographic distributions*. Princeton University Press, Princeton, NJ.
- Phillips, S.J. & Dudík, M. (2008) Modelling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- R Core Team (2014) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Radosavljevic, A. & Anderson, R.P. (2014) Making better MAXENT models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography*, **41**, 629–643.
- Randriamoria, T.M., Soarimalala, V. & Goodman, S.M. (2015) Terrestrial “forest-dwelling” endemic small mammals captured outside of natural habitats in the Moramanga District, central eastern Madagascar. *Malagasy Nature*, **9**, 97–106.
- Reddy, S. & Dávalos, L.M. (2003) Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, **30**, 1719–1727.
- Royle, J.A., Chandler, R.B., Yackulic, C. & Nichols, J.D. (2012) Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods in Ecology and Evolution*, **3**, 545–554.
- Shcheglovitova, M. & Anderson, R.P. (2013) Estimating optimal complexity for ecological niche models: a jackknife approach for species with small sample sizes. *Ecological Modelling*, **269**, 9–17.
- Soarimalala, V. & Goodman, S.M. (2011) *Les petits mammifères de Madagascar*. Association Vahatra, Antananarivo.
- Stephenson, P.J., Soarimalala, V. & Goodman, S. (2016a) *Microgale cowani*. The IUCN Red List of Threatened Species 2016: e.T40586A97190403. Available at: <http://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T40586A97190403>. en (accessed 26 August 2016).
- Stephenson, P.J., Soarimalala, V. & Goodman, S. (2016b) *Microgale gracilis*. The IUCN Red List of Threatened Species 2016: e.T13343A97191617. Available at: <http://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T13343A97191617>. en (accessed 26 August 2016).
- Stephenson, P.J., Soarimalala, V. & Goodman, S. (2016c) *Microgale gymnorhyncha*. The IUCN Red List of Threatened Species 2016: e.T62014A97191808. Available at: <http://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T62014A97191808>. en (accessed 26 August 2016).
- Stephenson, P.J., Soarimalala, V. & Goodman, S. (2016d) *Microgale jobihely*. The IUCN Red List of Threatened Species 2016: e.T136628A21282528. Available at: <http://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T136628A21282528>. en (accessed 26 August 2016).
- Stephenson, P.J., Soarimalala, V. & Goodman, S. (2016e) *Microgale thomasi*. The IUCN Red List of Threatened Species 2016: e.T13355A97202776. Available at: <http://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T13355A97202776>. en (accessed 26 August 2016).
- Thuiller, W., Georges, D. & Engler, R. (2014) Package ‘biomod2’: Ensemble Platform for Species Distribution Modelling. Available at: cran.r-project.org/web/packages/biomod2/.
- Veloz, S.D. (2009) Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, **36**, 2290–2299.
- Warren, D.L. & Seifert, S.N. (2011) Ecological niche modelling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**, 335–342.
- Warren, D.L., Glor, R.E., Turelli, M. & Funk, D. (2009) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868–2883, Erratum Evolution, 2865, 1215.
- Watling, J.I., Brandt, L.A., Bucklin, D.N., Fujisaki, I., Mazzotti, F.J., Románach, S.S. & Speroterra, C. (2015) Performance metrics and variance partitioning reveal sources of uncertainty in species distribution models. *Ecological Modelling*, **309**, 48–59.
- Wintle, B.A., McCarthy, M.A., Volinsky, C.T. & Kavanagh, R.P. (2003) The use of Bayesian model averaging to better represent uncertainty in ecological models. *Conservation Biology*, **17**, 1579–1590.
- Wis, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H. & Guisan, A. (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.

SUPPORTING INFORMATION

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

Figure S1. Results of threshold-dependent and threshold-independent evaluations in tuning experiments of MAXENT models.

Figure S2. Maps showing the three different ecological niche models for *Microgale gymnorhyncha*.

Figure S3. Maps showing the three different ecological niche models for *Microgale thomasi*.

Figure S4. Maps showing two different methods of displaying uncertainty for two species of tenrecs.

Figure S5. Maps for the four Malagasy shrew-tenrecs of the genus *Microgale* displaying the standard deviation of the continuous outputs.

Appendix S1. Contains the locality information and supplementary figure captions.

BIOSKETCH

Robert A. Boria began his research with niche modelling as an undergraduate at the City College of New York. For his Master's thesis, he used niche models to study small mammals in Madagascar. Currently, he is pursuing his PhD at

University of California, Merced conducting research on the evolutionary ecology of small mammals in North America.

Author contributions: R.A.B. and R.P.A. conceived the ideas; L.E.O. and S.M.G. provided locality data; R.A.B. and R.P.A. analysed the data; all authors interpreted results; R.A.B. and R.P.A. led the writing; and L.E.O. and S.M.G. contributed to writing.

Editor: Wilfried Thuiller