WILEY -Global Change Biology

Toward ecologically realistic predictions of species distributions: A cross-time example from tropical montane cloud forests

Lázaro Guevara[1,2] (iD) | Beth E. Gerstner[1,3] | Jamie M. Kass[1,4] | Robert P. Anderson[1,4,5]

[1]Department of Biology, City College of New York, City University of New York, New York, NY, USA

[2]Departamento de Biología Evolutiva, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City, Mexico

[3]Department of Fisheries & Wildlife, Michigan State University, East Lansing, MI, USA

[4]Program in Biology, Graduate Center, City University of New York, New York, NY, USA

[5]Division of Vertebrate Zoology (Mammalogy), American Museum of Natural History, New York, NY, USA

Correspondence
Lázaro Guevara, Department of Biology, City College of New York, City University of New York, New York, NY, USA.
Emails: llg@st.ib.unam.mx; guevalopez@yahoo.com.mx

Abstract

There is an urgent need for more ecologically realistic models for better predicting the effects of climate change on species' potential geographic distributions. Here we build ecological niche models using MAXENT and test whether selecting predictor variables based on biological knowledge and selecting ecologically realistic response curves can improve cross-time distributional predictions. We also evaluate how the method chosen for extrapolation into nonanalog conditions affects the prediction. We do so by estimating the potential distribution of a montane shrew (Mammalia, Soricidae, *Cryptotis mexicanus*) at present and the Last Glacial Maximum (LGM). Because it is tightly associated with cloud forests (with climatically determined upper and lower limits) whose distributional shifts are well characterized, this species provides clear expectations of plausible vs. implausible results. Response curves for the MAXENT model made using variables selected via biological justification were ecologically more realistic compared with those of the model made using many potential predictors. This strategy also led to much more plausible geographic predictions for upper and lower elevational limits of the species both for the present and during the LGM. By inspecting the modeled response curves, we also determined the most appropriate way to extrapolate into nonanalog environments, a previously overlooked factor in studies involving model transfer. This study provides intuitive context for recommendations that should promote more realistic ecological niche models for transfer across space and time.

KEYWORDS
climate change, ecological niche, Last Glacial Maximum, Mammalia, MAXENT, paleoecology, species distribution

1 | INTRODUCTION

Correlative ecological niche models (ENM; or species distribution models) characterize a species' environmental associations and estimate its potential geographical distribution (Peterson et al., 2011). These models can be transferred to estimate potential distributions in another place or time (i.e., future or past). Model transfer is increasingly used for forecasting species invasions and distributions under ongoing anthropogenic climate change, or hindcasting the impact of past climate change (Alvarado-Serrano & Knowles, 2014; Broennimann & Guisan, 2008; Elith, Kearney, & Phillips, 2010). Given the importance of such estimates of species potential distributions for biodiversity management and conservation as well as ecological and evolutionary studies, there is an urgent need for more ecologically realistic and transferable niche models (Moritz & Agudo, 2013; Owens et al., 2013).

Intuitively, if an ENM predicts an implausible current potential distribution, transfer to a different region or time period may lead to greater errors. Recommendations to yield more realistic and transferable niche models are not new (see Anderson, 2013; Araújo & Peterson, 2012; Austin, 2002, 2007; Austin & Van Niel, 2011; Elith & Leathwick, 2009; Rödder & Lötters, 2010), but they remain rarely implemented in empirical studies. For example, climatic variable predictors used to estimate species niches ideally should constitute driving factors that limit the species' distribution (Austin, 2007; Jiménez-Valverde et al., 2011; Kearney & Porter, 2009; Thuiller, Brotons, Araújo, & Lavorel, 2004). Nevertheless, a common practice is to include many plausible predictors (likely correlated with the unknown drivers), motivated more by the expectation that the algorithm will identify those that are ecologically important than by a well-founded biological justification. Failure to choose variables most related to the species' ecology can lead to implausible predictions of species responses to climate change (Austin, 2007; Morán-Ordóñez, Briscoe, & Wintle, 2017). In addition, besides choosing relevant predictor variables, there is a need for fitting response curves that are realistic based on ecological theory (Thuiller et al., 2004). Theory suggests that response shapes to climatic predictors should be smooth, unimodal, and roughly bell-shaped, though often skewed (Austin, 2002, 2007). However, visual inspection of the shape of the responses fitted by niche models to assess ecological realism is almost never undertaken.

A third critical reason for erroneous predictions upon model transfer is inappropriate methods for extrapolation of response curves into environmental conditions different than those existing in the region of model calibration (i.e., nonanalog conditions; Fitzpatrick & Hargrove, 2009). This situation arises when a species' response for a given variable is truncated (i.e., the new place or time period includes environmental conditions that are more extreme than those available for model calibration) and a prediction in another place or time is required (Fitzpatrick & Hargrove, 2009; Owens et al., 2013; Thuiller et al., 2004; Williams & Jackson, 2007). Using MAXENT, one of the most popular algorithms for niche modeling (Merow, Smith, & Silander, 2013), researchers have two options for predicting into nonanalog conditions: (i) allowing the model to estimate the response without any constraint; or (ii) holding ("clamping") the response constant at the modeled suitability value at the point of truncation (Elith et al., 2010; see also illustrations in Anderson, 2013). Visual inspection of response curves could allow detection of truncated responses for predictor variables and may prove useful as a way of choosing the most reasonable method for extrapolation (Anderson, 2013).

Here we assess how these ideas can influence the generation of better distributional predictions by estimating the cross-time potential distribution of a small mammal highly associated with tropical montane cloud forests. We specifically focus on its current potential distribution and model transfer into the Last Glacial Maximum (LGM), when colder and drier conditions existed. Despite the complications to validate potential distributions, realistic expectations for this mammal can be made based on its tight habitat affinity, both for the present and during the LGM, based on paleoecological data. Most of the issues affecting cross-time transfer of niche models are also relevant for transferring them across space, broadening the applicability of this study.

The cloud forest species considered in this study is the Mexican small-eared shrew (Soricidae, *Cryptotis mexicanus*) known to inhabit montane areas in central-eastern Mexico (Figure 1; Guevara, Cervantes, & Sánchez-Cordero, 2015). It is relatively well-sampled and common through its full geographic range (González-Ruiz, Ramírez-Pulido, & Gual-Díaz, 2014; Guevara et al., 2015). Cloud forests in northern Mesoamerica are associated with very humid areas and are usually distributed at mid-elevations below the frost line (Jardel-Peláez, Cuevas-Guzmán, Santiago-Pérez, & Rodríguez-Gómez, 2014). Consequently, cloud forests species tend to be physiologically constrained by cold and heat tolerance, as well as by moisture stress (Williams-Linera, 2007). Therefore, for such species, we expect that predictor variables capturing those extreme climatic conditions will likely be ideal to achieve realistic niche models. This natural history information leads to the following expectations:

1.1 | Expectation 1

Available information indicates that *C. mexicanus* currently can inhabit areas at middle-to-high elevations (but not extremely high ones)
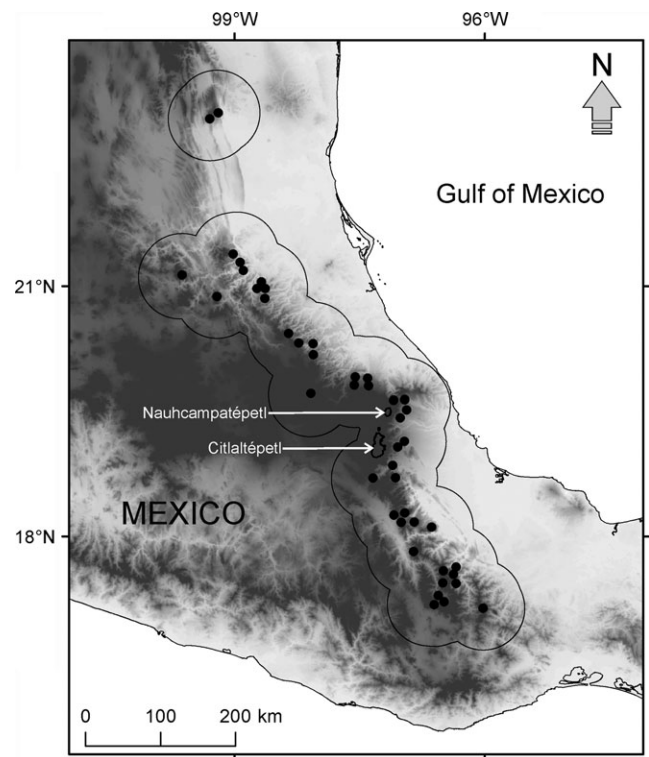

FIGURE 1 Map showing the occurrence data (black points), calibration region (buffer around occurrences), and projection region (full map) used to estimate the cross-time potential distribution of the Mexican small-eared shrew (*Cryptotis mexicanus*), in central-eastern Mexico. Names of main physiographic features mentioned in the text are included

if soil moisture is also high (Guevara et al., 2015). On the other hand, this species cannot inhabit lowlands due to high temperatures. This means that extremely cold conditions (which also tend to be dry) and warm ones (even if they are wet) should limit its potential distribution. Consequently, based on the current distribution of cloud forests in the region, a realistic niche model for this species in the present should predict areas above ≈3200 m and below ≈500 m elevation to be unsuitable.

1.2 | Expectation 2

Paleoecological data document that climate during the LGM in this region was colder and less rainy than today. Relative to present, temperatures during the LGM may have been 8.5°C degrees lower at elevations above ≈3,300 m in northern Mesoamerica, with many mountains covered by ice caps (Bush et al., 2009). Pollen and charcoal data for the LGM indicate that the upper limit of cloud forests was ≈1,000 m lower than at present (Caballero, Lozano-García, Vázquez-Selem, & Ortega, 2010; Correa-Metrio et al., 2012). Consequently, a realistic niche model for this species in the present projected to conditions for the LGM should predict areas above ≈2200 m elevation to be unsuitable.

Hence, in this study we test whether more realistic current and cross-time potential distributional estimates are obtained: (i) using predictor variables characterizing factors clearly associated with climatic limits to its distribution; and (ii) by selecting ecologically plausible response curves. In addition, we evaluate whether (iii) the extrapolation method chosen in MAXENT altered the prediction into nonanalog conditions.

2 | MATERIALS AND METHODS

2.1 | Input data

We used occurrence data from voucher specimens in natural history museums, with taxonomic identifications confirmed via morphology (Sánchez-Cordero & Guevara, 2016; see Table S1). Occurrence localities from specimen tag data were then thoroughly georeferenced (i.e., for those lacking field GPS readings; Sánchez-Cordero & Guevara, 2016). To reduce problems associated with spatial sampling biases (Boria, Olson, Goodman, & Anderson, 2014; Merow et al., 2013), we spatially thinned our original dataset consisting of 89 localities using the spThin package in R (Aiello-Lammens, Boria, Radosavljevic, Vilela, & Anderson, 2015). Thinning ensures that no two localities were within 10 km of one another, yielding 44 records in the final dataset. We ran spThin with 100 iterations, likely to be sufficient to retain the most localities possible (Aiello-Lammens et al., 2015).

As potential predictors of the species' niche, we used 19 bioclimatic variables from the WorldClim dataset (Table 1; see below for experimental design and selection of a subset of variables). These have a resolution of ca. 1 km^2 at the equator and reflect aspects of temperature and rainfall, and their seasonality (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). We delimited the area to calibrate

TABLE 1 Bioclimatic variables from the WorldClim dataset (Hijmans et al., 2005) commonly used to predict species distributions using correlative niche models

BIO 01	Annual mean temperature
BIO 02	Mean diurnal range
BIO 03	Isothermality
BIO 04	Temperature seasonality
BIO 05	Maximum temperature of warmest month
BIO 06	Minimum temperature of coldest month
BIO 07	Temperature annual range
BIO 08	Mean temperature of wettest quarter
BIO 09	Mean temperature of driest quarter
BIO 10	Mean temperature of warmest quarter
BIO 11	Mean temperature of coldest quarter
BIO 12	Annual precipitation
BIO 13	Precipitation of wettest month
BIO 14	Precipitation of driest month
BIO 15	Precipitation seasonality (coefficient of variation)
BIO 16	Precipitation of wettest quarter
BIO 17	Precipitation of driest quarter
BIO 18	Precipitation of warmest quarter
BIO 19	Precipitation of coldest quarter

The units for precipitation variables are mm (except for precipitation seasonality, which is the unitless coefficient of variation), whereas those for temperature variables are °C (except for isothermality, which is a unitless ratio, and temperature seasonality, which is standard deviation).

models by drawing a 0.5° circular buffer around the localities before thinning (Figure 1). This region includes environments that are probably accessible to the species given its dispersal limitations and the configuration of barriers (Anderson & Raza, 2010; Barve et al., 2011).

2.2 | Model evaluation and model selection

Given its frequent use to predict cross-time potential distributions, we built niche models using the presence-background technique MAXENT 3.3.3k (Phillips, Anderson, & Schapire, 2006). To select specific model settings approximating optimal levels of complexity, we made models with a wide variety of different combinations of feature classes (FC: Linear; Quadratic; Linear and Quadratic; Hinge; Linear, Quadratic, and Hinge) and regularization multipliers (RM: 1.0–5.0 with 0.5 intervals). This led to a total of 36 unique combinations of FCs and RMs. The optimal settings were selected based on the model with the lowest Akaike Information Criterion corrected for small sample sizes (AICc, Warren & Seifert, 2011) for all models of each experiment (see below). Model evaluation was performed in the ENMeval package in R (Muscarella et al., 2014). We applied a spatial block approach to data partitioning to yield model evaluation statistics. Spatial cross-validation is especially important when the aim is cross-time or cross-space model transfer (Roberts et al., 2017). We inspected threshold-dependent (omission rate for testing

points, or OR10, using a threshold set by the 10% training omission rate) and threshold-independent (AUC for testing points, or AUC_{TEST}) evaluation statistics to determine whether those models chosen as optimal by AICc are also high performing in terms of overfitting and discrimination, respectively (Anderson & Gonzalez, 2011).

2.3 | Final models and model transfer

Using the maxent.jar software version 3.3.3k, we made final models using all filtered localities, for the optimal settings identified above. Final models were projected onto present and LGM conditions for a bounding box encompassing the central-eastern region of Mexico (15°35′–22°44′N, 94°12′–102°35′W). We did this via the two extrapolation methods available in MAXENT: unconstrained extrapolation and extrapolation by "clamping." For the LGM, we employed estimates from the general circulation model produced by the Community Climate System Model (CCSM3; Collins et al., 2006), since it is the one available that appears to best describe the conditions during this time in northern Mesoamerica.

To illustrate the suitable areas within our final model predictions, we divided the continuous prediction into a binary one (suitable vs. unsuitable) using the lowest presence threshold (LPT) or minimum training presence threshold of MAXENT, where the training omission rate is equal to zero (Pearson, Raxworthy, Nakamura, & Peterson, 2007). The LPT sets the threshold at the lowest value of the prediction for any pixel that contains a training locality (Pearson et al., 2007). Suitability values above the LPT were preserved in the original continuous format.

2.4 | Experimental design

We conducted two sets of main experiments, each with two variants. First, we used the complete set of 19 bioclimatic variables (Table 1), which is a widely used strategy in current literature. Note, however, that if the RM is >0, the model may not incorporate the complete set of input variables, especially after estimating optimal levels of complexity. Second, we used four predictor variables selected with biological justification by expert opinion to capture extreme conditions, and thus more likely represent surrogates for factors limiting the distribution of the species. These four variables were: BIO 05, BIO 06, BIO 13, and BIO 14 (Table 1).

Finally, to evaluate the effect of potential artifactual truncation of response curves because of inadequate background sampling, we included two variants in our experiments. First, we used 10,000 randomly selected pixels within the study area as background in model calibration, which is the default and common practice for MAXENT. Second, we ensured a full representation of environments available for the species by including all pixels within the delimited study area (i.e., fully sampled background that contains 135,827 pixels, rather than that random sample of 10,000 pixels). In summary, we produced four suites of models: 1a) all 19 bioclimatic variables with either 10,000 random background pixels [BC19$_{rand}$], or 1b) fully sampled background [BC19$_{full}$]; 2a) four expert-selected bioclimatic variables

with 10,000 random background points [BC4$_{rand}$], or 2b) fully sampled background [BC4$_{full}$]).

For the two experiments, each with its two variants, we assessed final models for ecological realism with regard to our two expectations. We did so in both environmental and geographic spaces by examining the resulting response curves and geographic predictions, respectively. For the former, we evaluated ecological realism by visually inspecting the modeled responses with respect to each environmental variable incorporated in the final model (e.g., smooth and roughly bell-shaped; Austin, 2002, 2007). These plots also indicate any environmental truncations in the study area, as well as the shape of any extrapolated response into nonanalog conditions (via selection of unconstrained or clamped extrapolation; Phillips et al., 2006; Elith et al., 2010). Finally, we visually inspected projections of the models onto the study region mentioned above, at each of the time periods.

3 | RESULTS

3.1 | Model selection and evaluation

Using the AICc to select optimal settings yielded clear trends regarding FC, RM, and evaluation metrics (Table 2). Using 19 variables, the optimal model settings were the same, irrespective of the background used to calibrate the models (hinge feature class with a RM of 4.5). Evaluation statistics were almost identical for both of those models. Each had the highest AUC_{TEST} values (BC19$_{rand}$: 0.832, BC19$_{full}$: 0.834) and one of the lowest OR10s (0.068 for both) among the respective candidate models. Similarly, when using the four expert-selected predictor variables, the optimal model settings were also the same no matter the background used (Linear and Quadratic FC with RM 1.0). In this case, each AICc-selected model had the lowest average omission rate (0.023 for both) and one of the highest AUC_{TEST} values (BC19$_{rand}$: 0.833, BC19$_{full}$: 0.831) among the candidates.

3.2 | Examining variables used and response curves

Predictor variables that were included in the final models varied between the two experiments (Table S2). The BC19 models used

TABLE 2 Evaluation metrics of ecological niche models generated for the small-eared shrew, *Cryptotis mexicanus*, in central-eastern Mexico

Model	FC	RM	AUC$_{TEST}$	OR10	pm
BC19$_{rand}$	H	4.5	0.832	0.068	3
BC19$_{full}$	H	4.5	0.834	0.068	3
BC4$_{rand}$	LQ	1	0.833	0.023	6
BC4$_{rand}$	LQ	1	0.831	0.023	6

FC, feature classes (H = Hinge, L = Linear, Q = Quadratic); RM, regularization multiplier; AUC$_{TEST}$, the threshold-independent metric based on predicted values for the test localities; OR10, the threshold-dependent metric that indicates the proportion of test localities with suitability values lower than that which excludes the 10% of training localities with the lowest predicted suitability; pm, number of parameters in the model.

only three of the 19 variables initially provided: BIO 01, BIO16, and BIO 19. The BC4 models used all four input variables. Response curves fitted by BC4 models were more ecologically realistic than those of the BC19 models (Figure 2). For the BC4 models, response curves were all smooth and unimodal: bell-shaped for BIO 06 and BIO 14, whereas skewed for BIO 05 and BIO 13. For the BC19 model, response curves for the three variables incorporated in the final models resulted in irregularly shaped and more complex ecological responses, showing sudden declines.

Importantly, truncation points for the response curves differed dramatically depending on the environmental background sampling, and the manner of environmental extrapolation led to key differences in geographic predictions. The random sample of 10,000 pixels, representing 7.3% of the training study area, failed to capture all environmental conditions available for this species, particularly those poorly represented in geography at present (e.g., very cold temperatures at the mountaintops; Figure 3). This underrepresentation of environments resulted in an artifactual truncation of response curves

in the BC19$_{rand}$ and BC4$_{rand}$ models. In contrast for example, the BC4$_{rand}$ model that predicted the response into nonanalog conditions without constraints yielded smoother and more realistic response curves, as opposed to those clamped to the most-extreme value within the calibration sample.

3.3 | Current potential distributions

The BC19 and BC4 models of *C. mexicanus* differed substantially in geographic space, with the BC4 models yielding more ecologically realistic distributional predictions. In general, BC19 models predicted a larger and more diffuse potential distribution than the BC4 models (Figure 4). The BC19 models showed extensive highlands as suitable for this cloud forest species. The highest suitability scores indicated for these predictions largely corresponded to mountains; this even included several mountaintops harboring extremely cold habitats above 3200 m elevation, such as Citlaltépetl (i.e., Pico de Orizaba), which at 5636 m elevation is the highest volcano in Mexico and
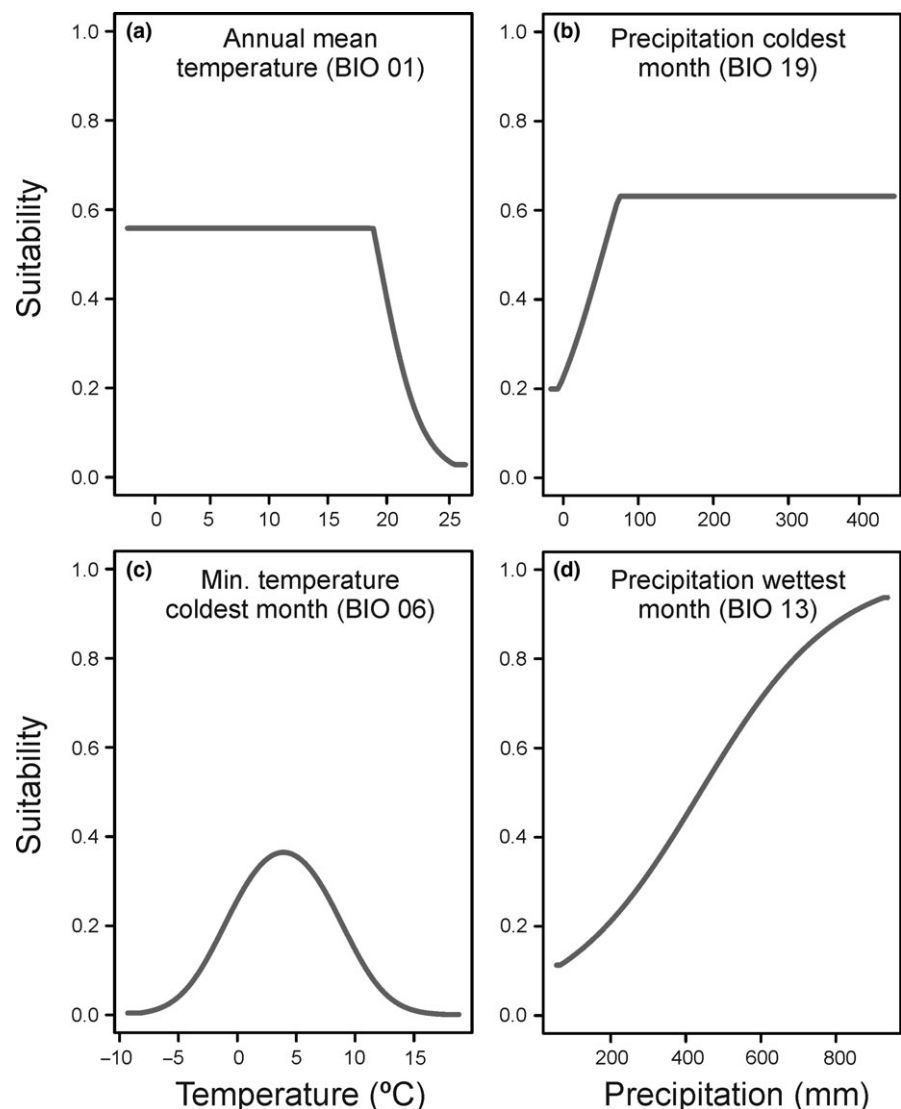

FIGURE 2 Response curves for two of the most important predictor variables for MAXENT models of the Mexican small-eared shrew (*Cryptotis mexicanus*), in central-eastern Mexico (Table S2). Response curves correspond to models built using 19 (a, b) and 4 variables (c, d). The y-axes show suitability values ranging from 0.0 (lowest) to 1.0 (highest). The x-axes for (a) and (b) are temperature in °C, and those for (b) and (d) are precipitation in mm
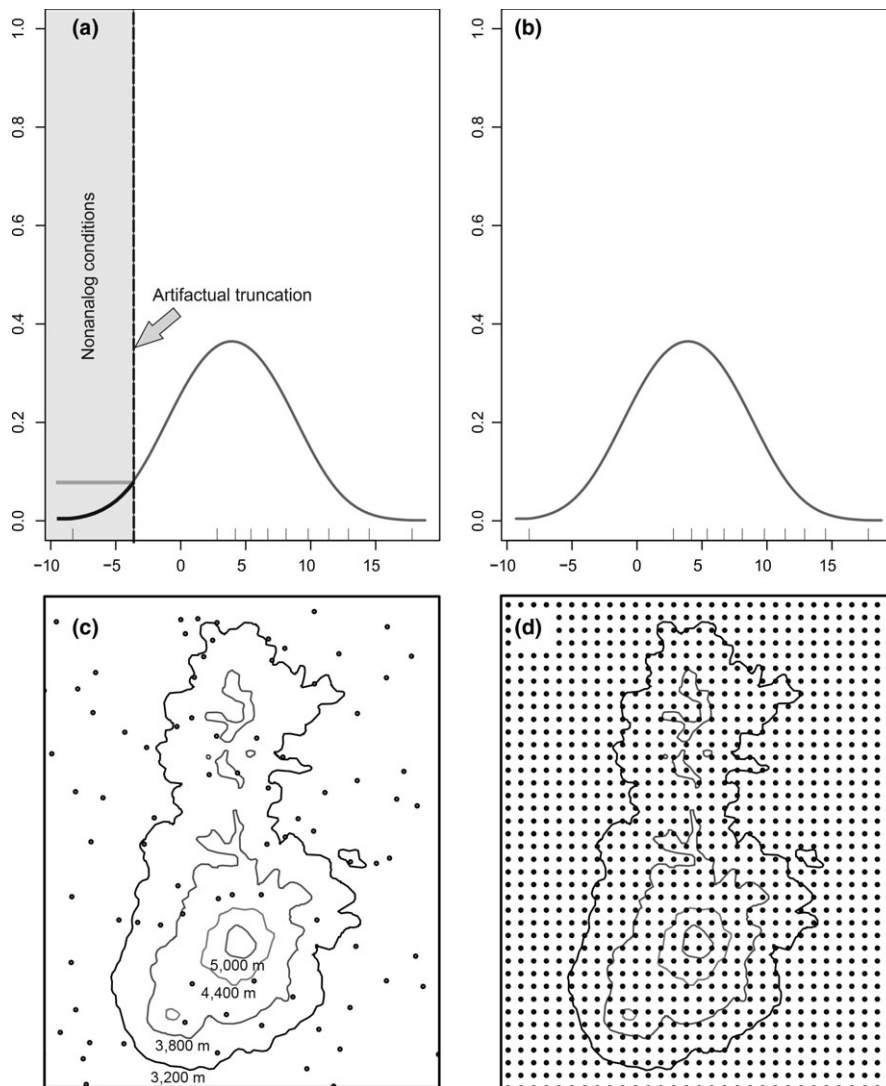
FIGURE 3 Response curves for the most important variable in the models (BIO 06, min temperature of coldest month in °C) based on permutation importance using four variables (Table S2), and zoom of background selection for a critical high-elevation region. Note slight differences in the shape of the respective response curves (a and b), depending on the number and location of pixels selected as background (c and d, close-up of the Citlaltépetl); however, a great difference exists in the environmental domain available for model calibration (left portion of a and b). Panel (a) shows the two methods for extrapolation into nonanalog conditions (gray line: clamping; black line: unconstrained), whereas (b) illustrates how a full representation of environments prevented the need for extrapolation (and hence, any associated complications) when projecting the model into present conditions in the study area. Such subtle differences can lead to much greater discrepancies in both environmental and geographic spaces when applied across space or time (e.g., to the Last Glacial Maximum)

holds permanent snowcaps. Surprisingly for a cloud forest species, areas with alpine tundra at Citlaltépetl were also predicted as suitable. The BC19 models also predicted moderately high suitability in some warm-wet lowland forests located below 500 m elevation.

In contrast, the BC4 models showed substantially greater discrimination (between the lowest and highest suitability values) and a tighter prediction that was less diffuse in geographic space. Notably, they indicated lower suitability in the highest alpine forests and icecaps near the summit of Citlaltépetl than the BC19 models. The BC4 models also showed a lower suitability in the lowlands, indicating barriers to dispersal and population connectivity among highland areas. In addition, the BC4 models showed greater suitability values in cloud forest areas, compared with only moderate (never extremely high) suitability according to BC19 models. In sum, the BC4 models better fit with Expectation 1, showing more realistic niche models.

The BC4$_{rand}$ showed interesting patterns of variation between the two methods for extrapolation into nonanalog conditions (Figure S1). Despite the overall high similarity between both geographic predictions, a notable difference exists at the summit of Citlaltépetl. When clamping was enforced, high suitability was indicated at the

top of Citlaltépetl, whereas it decreased when the response curve was extrapolated without constraint. Importantly, the issue about how to extrapolate into nonanalog conditions in the same time/place was avoided when the full area was used as background (because by definition extrapolation was not required within the same study area/time period).

3.4 | Model transfer into the LGM

When models were projected to the LGM (including extrapolation into substantially nonanalog conditions), estimates of suitability differed dramatically between the BC19 and the BC4 models. Again, the BC19 models predicted a considerably larger geographic extent of moderately suitable areas during the LGM than the BC4 models (Figure 5). Contrary to Expectation 2, many pixels above 2200 m elevation showed a moderate environmental suitability in the BC19 models. Most strikingly, they considered as suitable the whole Citlaltépetl, as well as another very high peak, Nauhcampatépetl (i.e., Cofre de Perote; with 4,282 m elevation), even though with colder temperatures glaciers had descended considerably in elevation at that time.
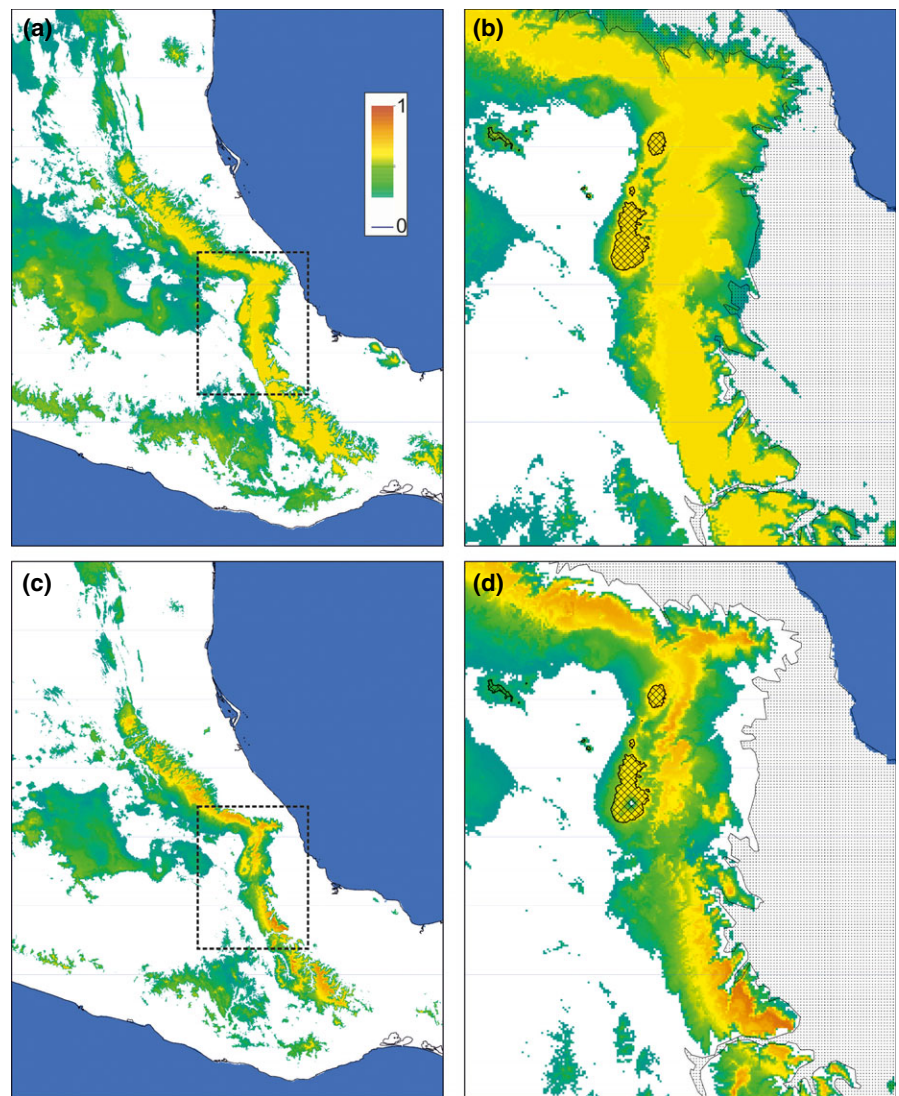
FIGURE 4 Current projections of Maxent models for the small-eared shrew, *Cryptotis mexicanus*, in central-eastern Mexico, using 19 (a) and 4 (c) variables in the models' construction. Letters (b) and (d) correspond to zooms for the predictions in the highest regions for each case. Cross-hatch covers the areas above 3200 m, whereas stippling shows areas below 500 m; both correspond to areas that should not hold suitable conditions for the species based on available natural history information (see Expectation 1 in main text)

In contrast, the BC4 models again led to more realistic inferences in the highlands, where high (and even moderate) suitability was assigned exclusively to areas below 2200 m, which matches Expectation 2 (Figure 5). However, this prediction was obtained consistently only when the full background was used for calibration (BC4$_{full}$). When we used the 10,000 random pixels (BC4$_{rand}$), the geographical prediction above 3200 m depended on the two different ways MAXENT deals with nonanalog conditions. Only the projection built with extrapolation without constraint showed low environmental suitability at those elevations (like the BC4$_{full}$), whereas clamping the response indicated a high (and implausible) suitability (Figure S1).

4 | DISCUSSION

4.1 | Ecological realism

This study demonstrates the benefit of choosing predictor variables based on a biological justification of the factors that are likely limiting the distribution of the focal species. These tests show that this strategy improved the prediction for current upper and lower elevational limits for a species inhabiting tropical montane cloud forest and for when the model was transferred to LGM climates (Figures 4 and 5). Unrealistic predictions were produced when we used the full set of 19 bioclimatic variables, especially at high elevations. Because the BC19 models integrated only three predictor variables, the results appear to be due to the nature of bioclimatic variables used (and/or how they are fitted), rather than to the number of variables incorporated into the model per se (Table 2).

Whenever available, physiological information from laboratory experiments should be taken into account for preselecting the variables used in correlative modeling of ecological niches (Kearney & Porter, 2009). Unfortunately, without a priori knowledge about the species' physiology, choice of variables thought to be causal or related to limiting factors is complicated (Guisan & Zimmermann, 2000; Petitpierre, Broennimann, Kueffer, Daehler, & Guisan, 2017). In such cases, investigators could consider natural history information about the species to select the most reasonable set of plausible environmental drivers of its distribution. The use of field notes or information about habitat affinity of the species can help for
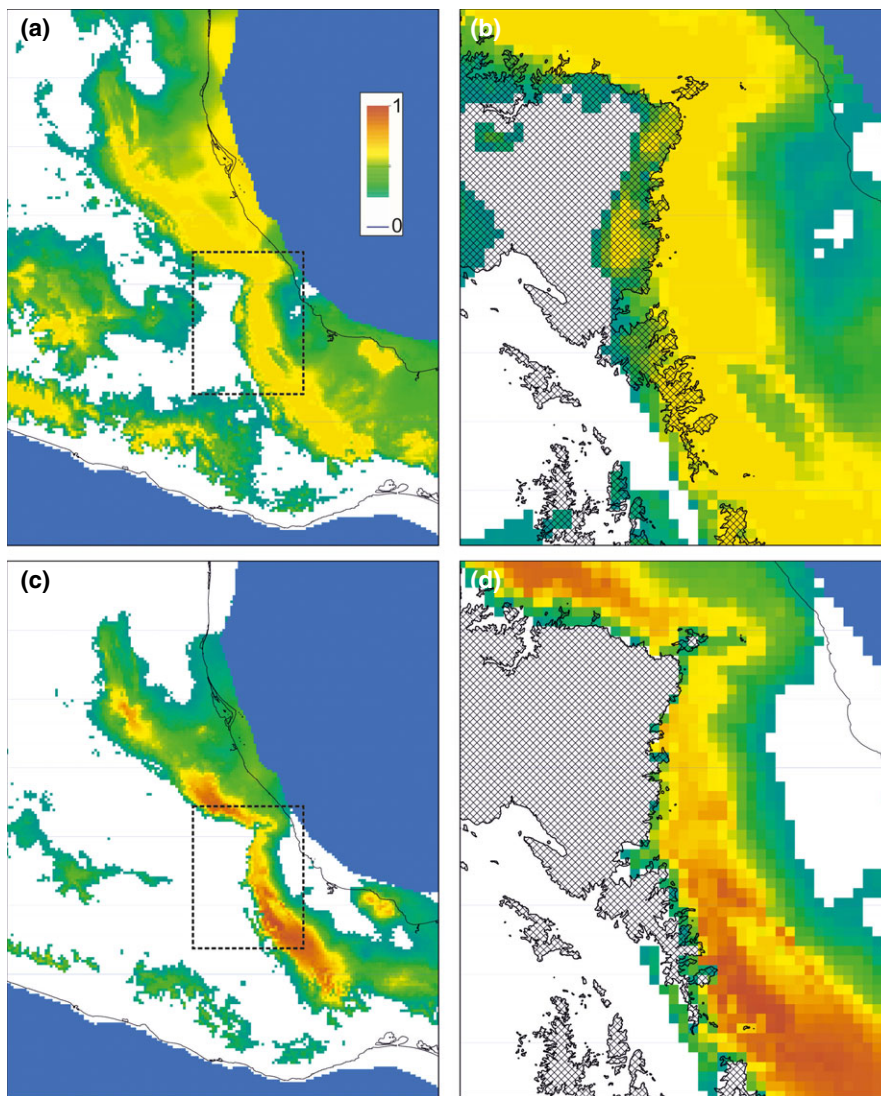
FIGURE 5 Reconstructed potential distribution during the Last Glacial Maximum (based on the CCSM scenario) for the small-eared shrew, *Cryptotis mexicanus*, in central-eastern Mexico, using 19 (a) and 4 (c) variables in the models' construction. Letters (b) and (d) correspond to zooms for the predictions in the highest regions for each case. Cross-hatch covers areas above 2200 m elevation, which approximates the upper limit of cloud forests at that time according to paleoecological information (see Expectation 2 in main text)

choosing the most justified predictors. For *C. mexicanus*, the choice of predictor variables likely to prove physiologically challenging for cloud forests species seems to be ecologically well-justified. Our findings suggest that extreme climatic conditions might limit species distributions instead of long-term mean climatic conditions (see Morán-Ordóñez et al., 2017).

This study also backs up theory indicating the value of models with ecologically realistic response curves to allow more plausible and transferable models (Owens et al., 2013). Ecologically implausible modeled response curves are difficult to interpret and problematic when the model is used to infer environmental correlates of species distributions (Araújo & Peterson, 2012; Elith & Leathwick, 2009). Here, acceptable evaluation statistics by themselves were insufficient for obtaining a realistic geographic prediction, and an inspection of ecological realism by examining responses curves fitted by the model was also needed to identify a plausible model (Austin, 2007; Thuiller et al., 2004). In addition, MAXENT models given a large number of arbitrarily selected variables might be good predictors of suitability for the time frame and extent of the study simply

because the algorithm found a good fit, and not because of any reasons that shed light on biological processes.

In addition, for some of the experiments, inspecting the response curves to detect truncation points for variables included in final models proved critical to determine the most appropriate way to extrapolate into nonanalog conditions (i.e., here by clamping or not). Underrepresentation of suitable environmental conditions (environmental truncation) can be caused by two nonmutually exclusive reasons. First, as presented in the Introduction, it can result from a limited range of environmental conditions existing in the study region (i.e., realized environmental space sensu Jackson & Overpeck, 2000). If Citlaltépetl and other high mountains did not exist in our study system, we would not have been able to characterize the species' full environmental response to temperature, detecting the decline in suitability in those coldest regions. In such cases, as has been noted frequently, it is impossible to characterize full responses because of the geographic constraint of environments available for the species (e.g., species with distributions ranging to the top of mountains).

Second, however, *artifactual* truncation can result from poor methodological representation of the environmental background. In this study, the coldest environments above 4400 m elevation were not represented in the background when we used 10,000 randomly selected pixels across the study area, the default implementation of Maxent (see Figure 3). This situation truncated the response curve unnecessarily and led to unrealistic geographical predictions for both the current and past. For the BC4$_{rand}$, the predictions into the coldest conditions varied markedly according to the extrapolation method chosen, clearly failing to produce a realistic prediction when clamping was enforced. Detecting and dealing with *artifactual* truncations has been almost completely ignored in the literature (but see Owens et al., 2013). To reduce the potential effect of inadequate representation of the environmental background, we strongly recommend an exhaustive sampling of the environments that are available for the species in the study area, reducing the chances of this type of artifactual truncation. With modern computers, we consider this is quite feasible for species with either localized or regional distributions like *C. mexicanus*.

This example also leads to a more nuanced view than that of a recent proposal regarding the method of extrapolation with respect to the direction of suitability at the point of truncation. Anderson (2013) pointed out that the riskiest extrapolations correspond to situations where the modeled suitability was increasing when truncated. In contrast, response curves truncated when suitability is decreasing yield lesser risk under extrapolation into nonanalog conditions. Although those tendencies still remain true, this study illustrates how the directionality of the response curve at the truncation point affects decisions regarding the most reasonable method of environmental extrapolation. For increasing curves, clamping the response might prove more reasonable than allowing an unconstrained extrapolation—at least if it is not necessary to extrapolate very far into nonanalog conditions; Figure 6a). However, as illustrated here for decreasing curves, unconstrained extrapolation likely represents the most reasonable option (Figure 6b).

According to knowledge regarding cloud forests species generally and *C. mexicanus* in particular, implausible predictions were produced when we used the full set of 19 bioclimatic variables, especially at high elevations. For example, Citlaltépetl (5,636 m) and Nauhcampatépetl (4,282 m) received a high environmental suitability estimate for the current climate, and both volcanoes were still predicted as highly suitable for the LGM, even though with colder temperatures glaciers had descended considerably in elevation at that time (Caballero et al., 2010).

4.2 | Recommendations and future directions

Improved methods to predict cross-time and cross-space potential distributions is needed in the face of current threats to biodiversity. As mentioned above, our findings not only apply for cross-time distributional predictions, either past or future, but also for projections to other spaces (e.g., studies of biological invasions, including species associated with zoonotic diseases). Here we emphasize three lessons to be considered in further studies, at least in studies with explanatory goals beyond geographic predictions from the patterns of observed data (Araújo & Peterson, 2012; Elith & Leathwick, 2009): (i) to the degree possible, predictor variable selection should be species-specific and guided by expert knowledge and biological plausibility (Austin, 2002, 2007; Austin & Van Niel, 2011; Rödder & Lötters, 2010; Zeng, Low, & Yeo, 2016), (ii) shapes of response curves fitted by the model should be inspected for ecological realism (Elith et al., 2010; Guisan & Zimmermann, 2000; Merow et al., 2013), and (iii)
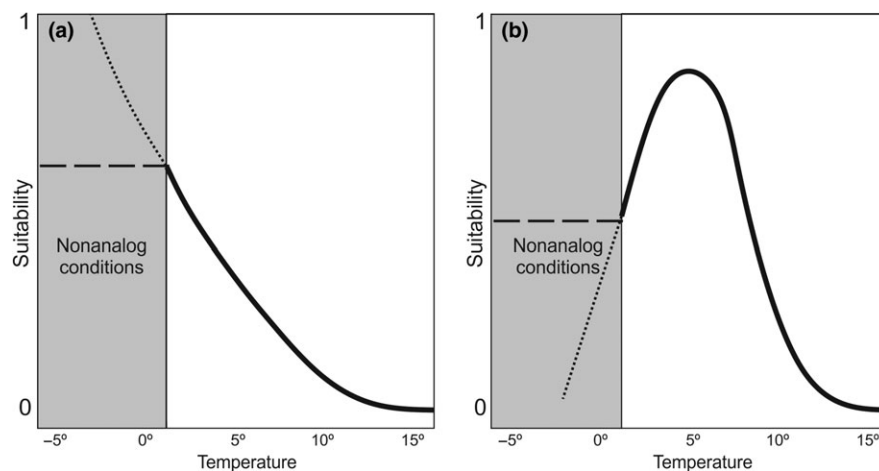

FIGURE 6 Hypothetical univariate response curves illustrating environmental truncations in the study region. Each panel illustrates a species' modeled response curve for a given climatic variable (*x*-axis), plotted against suitability (*y*-axis; 0–1). Transfer of a niche model to regions or time periods holding nonanalog conditions (gray area) requires extrapolation in environmental space. Two possibilities for extrapolation are shown: allowing the model to estimate the response without any constraint (dotted line) vs. "clamped" to the modeled suitability value at the point of truncation (dashed line). As mentioned in the text and based on the findings of this study, if a response curve is increasing at the point of truncation, clamping the response might prove more reasonable than allowing an unconstrained extrapolation (a). However, in cases of decreasing modeled suitability at the truncation point, unconstrained extrapolation likely represents the most reasonable option (b)

the methods for extrapolation into nonanalog conditions should be justified based on the truncation points identified in the response curves. Hopefully, the literature will soon accumulate a sufficient number of studies addressing these factors, allowing a synthesis and production of general recommendations.

Several additional paths should be addressed in future research in this and other systems. First, a different combination of predictor variables that are biologically plausible for our study system might also result in realistic estimations. Novel techniques accounting for the degree of causality and multicollinearity of different sets of predictor variables to build niche models should be encouraged (e.g., Petitpierre et al., 2017; Zeng et al., 2016). Second, reconstructed past potential distributions should be considered as hypotheses to be tested (Collevatti et al., 2013; Moreno-Amat et al., 2015). Current spatial patterns of genetic diversity and inferences regarding species demographic history can help test the realism of reconstructed/hypothesized paleodistributions (Carnaval, Hickerson, Haddad, Rodrigues, & Moritz, 2009; Ramírez-Barahona & Eguiarte, 2014). In addition to vegetational data from paleoecology (used here), whenever available, fossils should also be used to test hypothesized paleodistributions (Guralnick & Pearman, 2009; Martínez-Meyer, Peterson, & Hargrove, 2004). In our example of model transfer into the LGM, we only used one paleoclimatic scenario CCSM3, but a variety exist from different climate modeling groups (Varela, Lima-Ribeiro, & Terribile, 2015). Therefore, tests of paleodistributions should be conducted including more climatic scenarios to evaluate the variability and plausibility among them.

Third, the field requires easy to use tools to examine response curves. R packages or other software with such functionality are needed to visualize key aspects of response curves, such as artifactual truncation points, and novel environments beyond those points. Such advancements will allow researchers to decide how to extrapolate into nonanalog conditions, and modeling software also should allow researchers to enforce these decisions separately for each variable, and even for upper vs. lower truncations of a given variable (not presently possible with MAXENT). Together, these recommendations and future directions should allow for more ecologically realistic and useful niche models and better inferences across space and time, in particular regarding the effects of past and ongoing climate change on the geographic distributions of species.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

ORCID

Lázaro Guevara iD http://orcid.org/0000-0002-5485-7056

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. https://doi.org/10.1111/ecog.01132

Alvarado-Serrano, D. F., & Knowles, L. L. (2014). Ecological niche models in phylogeographic studies: Applications, advances and precautions. *Molecular Ecology Resources*, *14*, 233–248. https://doi.org/10.1111/1755-0998.12184

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. https://doi.org/10.1111/nyas.12264

Anderson, R. P., & Gonzalez, I. (2011). Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, *222*, 2796–2811. https://doi.org/10.1016/j.ecolmodel.2011.04.011

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. https://doi.org/10.1111/j.1365-2699.2010.02290.x

Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, *93*, 1527–1539. https://doi.org/10.1890/11-1930.1

Austin, M. P. (2002). Spatial prediction of species distribution: An interface between ecological theory and statistical modelling. *Ecological Modelling*, *157*, 101–118. https://doi.org/10.1016/S0304-3800(02)00205-3

Austin, M. (2007). Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, *200*, 1–19. https://doi.org/10.1016/j.ecolmodel.2006.07.005

Austin, M. P., & Van Niel, K. P. (2011). Improving species distribution models for climate change studies: Variable selection and scale. *Journal of Biogeography*, *38*, 1–8. https://doi.org/10.1111/j.1365-2699.2010.02416.x

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., ... Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, *222*, 1810–1819. https://doi.org/10.1016/j.ecolmodel.2011.02.011

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. https://doi.org/10.1016/j.ecolmodel.2013.12.012

Broennimann, O., & Guisan, A. (2008). Predicting current and future biological invasions: Both native and invaded ranges matter. *Biology Letters*, *4*, 585–589. https://doi.org/10.1098/rsbl.2008.0254

Bush, M. B., Correa-Metrio, A. Y., Hodell, D. A., Brenner, M., Anselmetti, F. S., Ariztegui, D., ... Gilli, A. (2009) Re-evaluation of climate change in lowland Central America during the Last Glacial Maximum using new sediment cores from Lake Petén Itzá, Guatemala. In F. Vimeux, F. Sylvestre & M. Khodri (Eds.), *Past Climate Variability in South America and Surrounding Regions* (pp. 113–128). Dordrecht, The Netherlands: Springer. https://doi.org/10.1007/978-90-481-2672-9

Caballero, M., Lozano-García, S., Vázquez-Selem, L., & Ortega, B. (2010). Evidencias de cambio climático y ambiental en registros glaciales y en cuencas lacustres del centro de México durante el último máximo glacial. *Boletín de la Sociedad Geológica Mexicana*, 62, 359–377.

Carnaval, A. C., Hickerson, M. J., Haddad, C. F., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, 323, 785–789. https://doi.org/10.1126/science.1166955

Collevatti, R. G., Terribile, L. C., Oliveira, G., Lima-Ribeiro, M. S., Nabout, J. C., Rangel, T. F., & Diniz-Filho, J. A. F. (2013). Drawbacks to palaeodistribution modelling: The case of South American seasonally dry forests. *Journal of Biogeography*, 40, 345–358. https://doi.org/10.1111/jbi.12005

Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., ... Smith, R. D. (2006). The community climate system model version 3 (CCSM3). *Journal of Climate*, 19, 2122–2143. https://doi.org/10.1175/JCLI3761.1

Correa-Metrio, A., Bush, M. B., Cabrera, K. R., Sully, S., Brenner, M., Hodell, D. A., ... Guilderson, T. (2012). Rapid climate change and no-analog vegetation in lowland Central America during the last 86,000 years. *Quaternary Science Reviews*, 38, 63–75. https://doi.org/10.1016/j.quascirev 2012.01.025

Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1, 330–342. https://doi.org/10.1111/j.2041-210X.2010.00036.x

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Fitzpatrick, M. C., & Hargrove, W. W. (2009). The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation*, 18, 2255–2261. https://doi.org/10.1007/s10531-009-9584-8

González-Ruiz, N., Ramírez-Pulido, J., & Gual-Díaz, M. (2014). Mamíferos del bosque mesófilo de montaña en México. In M. Gual-Díaz & A. Rendón-Correa (Eds.), *Bosques mesófilos de montaña de México: Diversidad, ecología y manejo* (pp. 305–326). Mexico City, Mexico: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.

Guevara, L., Cervantes, F. A., & Sánchez-Cordero, V. (2015). Riqueza, distribución y conservación de los topos y las musarañas (Mammalia, Eulipotyphla) de México. *Therya*, 6, 43–68. https://doi.org/10.12933/therya-15-211

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Guralnick, R., & Pearman, P. B. (2009). *Using species occurrence databases to determine niche dynamics of montane and lowland species since the Last Glacial Maximum. Data mining for global trends in mountain biodiversity* (pp. 125–134). Boca Raton, FL: Taylor & Francis Group. http://dx.doi.org/10.1201/9781420083705.ch13

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. https://doi.org/10.1002/joc.1276

Jackson, S. T., & Overpeck, J. T. (2000). Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, 26, 194–220. https://doi.org/10.1666/0094-8373(2000)26[194:ROPPAC]2.0.CO;2

Jardel-Peláez, E., Cuevas-Guzmán, R., Santiago-Pérez, A. L., & Rodríguez-Gómez, J. M. (2014). Ecología y manejo de los bosques mesófilos de montaña en México. In M. Gual-Díaz & A. Rendón-Correa (Eds.), *Bosques mesófilos de montaña de México: Diversidad, ecología y manejo* (pp. 141–187). Mexico City, Mexico: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.

Jiménez-Valverde, A., Peterson, A. T., Soberón, J., Overton, J. M., Aragón, P., & Lobo, J. M. (2011). Use of niche models in invasive species risk assessments. *Biological Invasions*, 13, 2785–2797. https://doi.org/10.1007/s10530-011-9963-4

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. *Ecology Letters*, 12, 334–350. https://doi.org/10.1111/j.1461-0248.2008.01277.x

Martínez-Meyer, E., Peterson, A. T., & Hargrove, W. W. (2004). Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Global Ecology and Biogeography*, 13, 305–314. https://doi.org/10.1111/j.1466-822X.2004.00107.x

Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., McMahon, S. M., Normand, S., ... Elith, J. (2014). What do we gain from simplicity versus complexity in species distribution models?. *Ecography*, 37, 12671281. https://doi.org/10.1111/ecog.00845

Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, 36, 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Morán-Ordóñez, A., Briscoe, N. J., & Wintle, B. A. (2017). Modelling species responses to extreme weather provides new insights into constraints on range and likely climate change impacts for Australian mammals. *Ecography*, 40, 001–013. https://doi.org/10.1111/ecog.02850

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. https://doi.org/10.1016/j.ecolmodel.2015.05.035

Moritz, C., & Agudo, R. (2013). The future of species under climate change: Resilience or decline? *Science*, 341, 504–508. https://doi.org/10.1126/science.1237190

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. https://doi.org/10.1111/2041-210X.12261

Owens, H. L., Campbell, L. P., Dornak, L. L., Saupe, E. E., Barve, N., Soberón, J., ... PetersonA. T. (2013). Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling*, 263, 10–18. https://doi.org/10.1016/j.ecolmodel.2013.04.011

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. https://doi.org/10.1111/j.1365-2699.2006.01594.x

Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). *Ecological niches and geographic distributions. Monographs in Population Biology*, 49. Princeton, NJ: Princeton University Press.

Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C., & Guisan, A. (2017). Selecting predictors to maximize the transferability of species distribution models: Lessons from cross-continental plant invasions. *Global Ecology and Biogeography*, 26, 275–287. https://doi.org/10.1111/geb.12530

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Ramírez-Barahona, S., & Eguiarte, L. E. (2014). Changes in the distribution of cloud forests during the last glacial predict the patterns of genetic diversity and demographic history of the tree fern *Alsophila firma* (Cyatheaceae). *Journal of Biogeography*, *41*, 2396–2407. https://doi.org/10.1111/jbi.12396

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., ... Doorman, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*, 913–929. https://doi.org/10.1111/ecog.02881

Rödder, D., & Lötters, S. (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. https://doi.org/10.1007/s00114-010-0694-7

Sánchez-Cordero, V., & Guevara, L. (2016). *Modelado de la distribución potencial de las musarañas (Mammalia, Soricidae)*. Mexico City, Mexico: Universidad Nacional Autónoma de México. Instituto de Biología. Databases SNIB-CONABIO, project JM044.

Thuiller, W., Brotons, L., Araújo, M. B., & Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, *27*, 165–172. https://doi.org/10.1111/j.0906-7590.2004.03673.x

Varela, S., Lima-Ribeiro, M. S., & Terribile, L. C. (2015). A short guide to the climatic variables of the Last Glacial Maximum for biogeographers. *PLoS One*, *10*, e0129037. https://doi.org/10.1371/journal.pone.0129037

Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecological Applications*, *21*, 335–342. https://doi.org/10.1890/10-1171.1

Williams, J. W., & Jackson, S. T. (2007). Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, *5*, 475–482. https://doi.org/10.1890/070037

Williams-Linera, G. (2007). *El bosque de niebla del centro de Veracruz: Ecología, historia y destino en tiempos de fragmentación y Cambio Climático*. Xalapa, Mexico: Instituto de Ecología, A. C. & CONABIO.

Zeng, Y., Low, B. W., & Yeo, D. C. (2016). Novel methods to select environmental variables in MaxEnt: A case study using invasive crayfish. *Ecological Modelling*, *341*, 5–13. https://doi.org/10.1016/j.ecolmodel.2016.09.019

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.