A retrospective approach for evaluating ecological niche modeling transferability over time: the case of Mexican endemic rodents

Claudia N. Moreno-Arzate [1, 2] , **Enrique Martínez-Meyer** Corresp. [3, 4]

[1] Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico
[2] Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, Mexico City, Mexico
[3] Departamento de Zoología, Instituto de Biología,, Universidad Nacional Autónoma de México, Mexico City, Mexico
[4] Laboratorio Nacional Conahcyt sobre la Biología del Cambio Climático, Mexico City, Mexico

Corresponding Author: Enrique Martínez-Meyer
Email address: emm@ib.unam.mx

Ecological niche modeling (ENM) is a valuable tool for inferring suitable environmental conditions and estimating species' geographic distributions. ENM is widely used to assess the potential effects of climate change on species distributions; however, the choice of modeling algorithm introduces substantial uncertainty, especially since future projections cannot be properly validated. In this study, we evaluated the performance of seven popular modeling algorithms—Bioclim, Generalized Additive Models (GAM), Generalized Linear Models (GLM), Boosted Regression Trees (BRT), Maxent, Random Forest (RF), and Support Vector Machine (SVM)—in transferring ENM across time, using Mexican endemic rodents as a model system. We used a retrospective approach, transferring models from the near past (1950-1979) to more recent conditions (1980-2009) and vice versa, to evaluate their performance in both forecasting and hindcasting. Consistent with previous studies, our results highlight that input data quality and algorithm choice significantly impact model accuracy, but most importantly, we found that algorithm performance varied between forecasting and hindcasting. While no single algorithm outperformed the others in both temporal directions, RF generally showed better performance for forecasting, while Maxent performed better in hindcasting, though it was more sensitive to small sample sizes. Bioclim consistently showed the lowest performance. These findings underscore that not all species or algorithms are suited for temporal projections. Therefore, we recommend testing multiple algorithms using a retrospective approach before applying models to future scenarios.

A retrospective approach for evaluating ecological niche modeling transferability over time: the case of Mexican endemic rodents

Claudia Noemi Moreno-Arzate[1,2], Enrique Martínez-Meyer[3,4]

[1] Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, Mexico City, Mexico
[2] Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico
[3] Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, Mexico
[4] Laboratorio Nacional Conahcyt sobre la Biología del Cambio Climático, Mexico City, Mexico

Corresponding Author:
Enrique Martínez-Meyer[2]
Circuito Exterior S/N, Ciudad Universitaria, Mexico City, 04510, México
Email address: emm@ib.unam.mx

Abstract

Ecological niche modeling (ENM) is a valuable tool for inferring suitable environmental conditions and estimating species' geographic distributions. ENM is widely used to assess the potential effects of climate change on species distributions; however, the choice of modeling algorithm introduces substantial uncertainty, especially since future projections cannot be properly validated. In this study, we evaluated the performance of seven popular modeling algorithms—Bioclim, Generalized Additive Models (GAM), Generalized Linear Models (GLM), Boosted Regression Trees (BRT), Maxent, Random Forest (RF), and Support Vector Machine (SVM)—in transferring ENM across time, using Mexican endemic rodents as a model system. We used a retrospective approach, transferring models from the near past (1950-1979) to more recent conditions (1980-2009) and vice versa, to evaluate their performance in both forecasting and hindcasting. Consistent with previous studies, our results highlight that input data quality and algorithm choice significantly impact model accuracy, but most importantly, we found that algorithm performance varied between forecasting and hindcasting. While no single algorithm outperformed the others in both temporal directions, RF generally showed better performance for forecasting, while Maxent performed better in hindcasting, though it was more sensitive to small sample sizes. Bioclim consistently showed the lowest performance. These findings underscore that not all species or algorithms are suited for temporal projections. Therefore, we recommend testing multiple algorithms using a retrospective approach before applying models to future scenarios.

Introduction

Climate change is significantly impacting biodiversity, causing shifts in species abundance and distribution that lead to extensive reshuffling of biotas (Parmesan & Yohe, 2003; Pacifici et al., 2015; Peterson et al., 2015; MacLean et al., 2018; Widick & Bean, 2019; Habibullah et al., 2022; Jaroszynska et al., 2023). In response to these dynamic changes, ecological niche modeling (ENM) has emerged as a valuable tool for analyzing and predicting species' geographic distributions under various climatic scenarios, both past and future. Grounded in ecological niche theory, ENM integrates methodologies from multiple disciplines, including informatics, geographic information systems (GIS), and statistics (Austin, 2007; Soberón & Nakamura, 2009; Wiens et al., 2009; Sillero et al., 2021). This correlative approach utilizes georeferenced species occurrence data and environmental predictor variables to model the ecological niche of species, and projecting this information onto geographic space to generate a map that is commonly referred to as the species' potential distribution (Barbet-Massin & Jetz, 2014; Elith et al., 2006; Peterson, 2011; Soley-Guardia, Alvaro-Serrano & Anderson, 2024).

Several methods for constructing niche models have evolved to address different types of occurrence data: presence-only (e.g., Bioclim, environmental distances), presence-absence (e.g., Generalized Linear Models [GLM], Generalized Additive Models [GAM]), presence-pseudoabsence (e.g., Genetic Algorithm for Rule-set Prediction [GARP]), and presence-background (e.g., Ecological Niche Factor Analysis [ENFA], Maxent) (Phillips, 2008; Barbet-Massin et al., 2012; Barbet-Massin & Jetz, 2014; Warton & Aarts, 2013; Fan et al., 2018; Qiao et al., 2019; Sillero et al., 2023). Presence-absence methods are considered more robust when unequivocal absence data are available, as they help identify unsuitable areas that might be misclassified by presence-only methods (Brotons et al., 2004; Golicher et al., 2012). However, reliable absence data are rare, leading to the predominance of presence-only, presence-pseudoabsence, and presence-background algorithms (Vaz, Cunha & Nabout, 2015; Soley-Guardia, Alvaro-Serrano & Anderson, 2024; Sillero et al., 2021).

ENM is extensively used to evaluate the potential impacts of climatic changes by projecting modeled niches under current conditions onto past or future scenarios, a concept known as model transferability (Pearson & Dawson, 2003; Peterson, Martínez-Meyer & González-Salazar, 2004; Thomas et al., 2004; Waltari & Guralnick, 2009; Heikkinen, Marmion & Luoto, 2012; Garcia et al., 2016; Zhu, Fan & Peterson, 2021; Sillero et al., 2021). The success of this transferability trlies on three critical assumptions: (1) the ecological niche remains stable during the period of transfer (Soberón & Nakamura, 2009); (2) the relationship between environmental variables and species remains constant during climatic changes (Hijmans & Graham, 2006; Wiens et al., 2009); and (3) the species is in equilibrium with the environment in the calibration scenario, meaning they occupy the available suitable areas accessible to them (Sequeira et al., 2018; Yates et al., 2018). These premises are critical as they assume that the current relationship between the species and the environment is optimal, and transferability will be defined by how closely the model conforms to this relationship.

A major challenge in model transferability is the presence of non-analog climates—environmental conditions in the projected climatic scenario that are absent in the calibration scenario (Sequeira et al., 2018). Algorithms respond idiosyncratically to these conditions due to their programmatic features, frequently producing disparate results (Pearson, 2006; Araújo & Peterson, 2012; Essl et al., 2023). This issue is particularly problematic when projecting ecological niches into future climates, where no empirical data are available to validate algorithm performance. Consequently, a common approach for evaluating algorithm effectiveness in climate change studies involves projecting models between two historical periods for which data are available, such as from the recent past to the present (Rubidge et al., 2011; Piirainen et al., 2023).

Extensive research has been conducted to compare algorithm performance across spatial and temporal transferences, utilizing both real and virtual species (Prasad, Iverson & Liaw, 2006; Hijmans & Graham, 2006; Kharouba, Algar & Kerr, 2009; Dobrowski et al., 2011; Rubidge et al., 2011; Moreno-Amat et al., 2015; García-Callejas & Araújo, 2016) These studies reveal that results can vary greatly due to factors such as species traits, biotic interactions, data completeness, and climatic dissimilarities. Such variations complicate the identification of the specific impacts that algorithm choices have on the outcomes (Yates et al., 2018; Merow et al., 2014).

In this study, we evaluated the performance of seven popular modeling algorithms—Bioclim, Generalized Additive Models (GAM), Generalized Linear Models (GLM), Boosted

Regression Trees (BRT), Maxent, Random Forest (RF), and Support Vector Machine (SVM)—in transferring niche models of Mexican endemic rodents from the mid-20th century to the late-20th/early 21st centuries and vice versa. We selected Mexican endemic rodents due to their restricted and relatively well-known distributions, and because we do not expect significant climatic niche evolution over this period (Martínez-Meyer, Peterson & Hargrove, 2004). We expected that species with poor data quality (e.g., scarce or biased) would produce poor results regardless of the algorithm used. Conversely, for well-sampled species, we hypothesized that algorithms capable of modeling biologically meaningful response curves (e.g., bell-shaped responses to temperature), such as Maxent or SVM, would outperform simpler algorithms like Bioclim or GLM.

Materials & Methods

Occurrence data of species. We compiled occurrence records for 117 Mexican endemic rodent species (Ramírez-Pulido et al., 2014; Supplementary Material, Table S1) from various sources including natural history collections, journal articles, books, and theses (Supplementary Material, Table S2). Records span two periods, 1950-1979 and 1980-2009, to align with available climatologies for Mexico. We eliminated duplicate records and those with questionable taxonomic or geographic certainty, retaining species with at least 10 unique localities per period to reduce the risk of data incompleteness issues (Hernández et al., 2006). For algorithms requiring absence data, we generated pseudoabsences by randomly selecting non-presence localities matching the number of presence records using the Ecospat package in R 3.5, (Di Cola et al., 2017; R Core Team, 2017).

Climatic variables. We used 19 bioclimatic variables generated for Mexico for the mid-20th century (Time 1 [T1]: 1950-1979) and the late 20th/early 21st centuries (Time 2 [T2]: 1980-2009) (Cuervo-Robayo et al., 2020). These variables were derived from monthly averages of precipitation and minimum and maximum temperatures recorded at climate stations across Mexico, southern USA, northern Guatemala, and Belize. The resulting surfaces have a spatial resolution of 30 arc seconds (~1 km), following the methodology of the WorldClim dataset (Hijmans et al., 2005), and summarize the extreme, mean, and seasonal patterns in temperature and rainfall. To reduce model complexity and minimize overfitting, we conducted Pearson correlation analyses for each species and excluded variables with correlations above 0.80 (Radosavljevic & Anderson, 2014; Moreno-Amat et al., 2015; García-Callejas & Araújo, 2016; Regos et al., 2019) (Supplementary Material, Table S3). The area of analysis for each species (i.e., “M” in the BAM framework; Barve et al., 2011) was delineated by clipping the raster layers to the ecoregions (Olson et al., 2001) where each species has been recorded, assuming ecoregion boundaries serve as dispersal barriers (Radosavljevic & Anderson, 2014).

Ecological niche modeling. We evaluated the transfer capacity of seven algorithms: Bioclim (Nix & Busby, 1986; Beaumont, Hughes & Poulsen, 2005; Booth et al., 2014), a climatic envelope method; two regression-based techniques, Generalized Additive Models (GAM) and Generalized Linear Models (GLM) (Guisan, Edwards & Hastie, 2002); and four machine-learning algorithms: Boosted Regression Trees (BRT; Elith, Leathwick & Hastie, 2008), Maxent (Elith et al., 2006; Phillips, Anderson & Schapire, 2006; Merow, Smith & Silander, 2013), Random Forest (RF; Breiman, 2001; Prasad, Iverson & Liaw, 2006), and Support Vector Machine (SVM; Drake, Randin & Guisan, 2006). Detailed information about the functioning of each algorithm can be found in the references cited.

For implementation, we created models using the following R packages: *DISMO* for Bioclim (Hijmans et al., 2017), *SDM* for GLM, GAM, BRT, RF, and SVM. For GLM, we used a binomial response with a logit link, a quadratic function, and the Akaike Information Criterion (AIC) for stepwise selection (Guisan, Edwards & Hastie, 2002). We implemented GAM with a binomial response and a logit link function (Guisan, Edwards & Hastie, 2002). BRT was parameterized with a learning rate of 0.005, a tree complexity of 5, and a bag fraction of 0.5 (Elith, Leathwick & Hastie, 2008). RF was calibrated with 500 trees (Prasad, Iverson & Liaw, 2006). We used *ENMeval* (Kass et al., 2021) for Maxent models, testing four regularization multiplier values (0.5, 1, 1.5, 2) and combinations of five feature classes (linear, quadratic, product, threshold, and hinge), with clamping and extrapolation options disabled. The best model for each species was selected using the Akaike Information Criterion corrected for small sample sizes (AICc) (Warren et al., 2014). All output maps were expressed on a continuous scale from 0 to 1.

We allocated 70% of the occurrence records for model calibration and the remaining 30% for validation, applying the same proportions for pseudoabsences where required. The resulting models were then converted into binary maps (presence-absence) using a ten-percentile threshold to minimize overprediction from potentially erroneous data (Radosavljevic & Anderson, 2014). We evaluated the models within time periods using binomial tests that compared the results against random expectations (Anderson, Lew & Peterson, 2003).

Model transferences. The logical procedure to evaluate the capacity of algorithms to transfer niche models across different temporal scenarios involves calibrating a niche model in period 1 and transferring it onto the climatic scenario of period 2, then validating the transference with occurrences from period 2, or calibrating a niche model in period 2 and comparing the two maps (Hijmans & Graham, 2006). However, disparities in the number or environmental distribution of occurrences between periods can affect observed differences between the resulting maps, making it difficult to attribute such differences to algorithm performance alone. To address this issue, we implemented a cross-temporal approach to identify species with similar occurrence data structures in both time periods.

First, we calibrated a model using occurrences and climatic layers from period 1 ("auto1"); then, we generated a second model using the climatic layers from period 1 with occurrences from period 2 ("cross1"). Next, we calibrated a model using occurrences from period 2 with the climatic surfaces from that period ("auto2") and another model using occurrences from period 1 with the climatic surfaces from period 2 ("cross2"). All resulting maps were converted into binary format, and we compared "auto1" with "cross1" and "auto2" with "cross2" geographically (see next section). If the overall similarity between the two pairs of maps was less than 70%, it indicated significant differences in the number or distribution of occurrences between the time periods, which could hinder the ability of algorithms to transfer models across temporal scenarios. We repeated this procedure for all species and algorithms. Species for which the similarity value was below 70% were excluded from further analyses. Species with suitable datasets for analysis were classified as "control" species, while those without were labeled "without control." For control species, we proceeded with transferring models from period 1 to period 2 (forecast) and vice versa (hindcast) and subsequently evaluated model performance (Fig. 1).

Evaluation of algorithm performance. For each algorithm, models transferred from period 1 to period 2 were compared pixel-by-pixel against models calibrated for period 2, and vice versa. We constructed confusion matrices for these comparisons, using the calibration models as references. In the confusion matrix: "a" represents the number of presence pixels correctly predicted by the transferred model (sensitivity), "b" denotes the number of absence pixels incorrectly classified as presence (commission error or false positives), "c" indicates the number of presence pixels incorrectly classified as absence (omission error or false negatives), and "d" corresponds to the number of absence pixels correctly predicted (specificity). We then calculated the following indices to assess model performance: True Skill Statistics (TSS; Equation 1), Overlap Index (OI; Equation 2), False Negative Rate (FNR; Equation 3), and False Positive Rate (FPR; Equation 4) (Fielding & Bell, 1997):

$$TSS=\frac{a}{a+c}+\frac{d}{b+d}-1 \quad \text{Eq. 1}$$

$$OI=\frac{a}{a+c} \quad \text{Eq. 2}$$

$$FNR=\frac{b}{b+d} \quad \text{Eq. 3}$$

$$FPR=\frac{c}{a+c} \quad \text{Eq. 4}$$

TSS measures the accuracy of predictions by comparing the number of correctly predicted pixels to what would be expected by chance. It ranges from -1 (no better than random) to 1 (perfect discrimination), with values above 0.7 considered reliable (Allouche, Tsoar & Kadmon, 2006). OI measures the proportion of overlap between the maps from different time periods indicating their consistency. FNR represents the omission error and measures the rate of overfitting, with values ranging from 0 to 1, and FPR measures the overestimation, also ranging from 0 to 1 (Rebelo, Tarroso & Jones, 2010).

Statistical analysis. To evaluate differences between forecast and hindcast model transfers, we conducted a Mann-Whitney-Wilcoxon test (Pohlert, 2016). We also used a Kruskal-Wallis test to compare the performance of different algorithms in transferring models. When a significant difference was detected, we used a Nemenyi test for pairwise multiple comparisons of mean ranks among algorithms. These statistical tests were performed using the *PMCMR* package in R (Pohlert, 2016). Additionally, we explored the relationship between the number of occurrence records and model performance—measured by True Skill Statistics (TSS)—for each algorithm and direction (forecast and hindcast) using Pearson's correlation analyses. All calculations and statistical analyses were carried out in R 3.5 (R Core Team, 2017).

Results

Occurrence data of species. We compiled occurrence data for 117 Mexican endemic rodent species (Ramírez-Pulido et al., 2014). Of these, only 44 species had sufficient unique records (at least 10 per time period) to generate robust models (Supplementary Material, S3). Fourteen species were underrepresented in one of the two time periods, while 59 species lacked the minimum of 10 unique records in either period. Among the 44 species with sufficient data, *Peromyscus melanophrys* (n = 504), *Peromyscus difficilis* (n = 440), *Chaetodipus arenarius* (n = 248), and *Sigmodon mascotensis* (n = 191) had the highest number of records.

Temporal consistency of occurrence data. We evaluated the temporal consistency of occurrence data for the 44 species with sufficient records using a cross-validation test. None of the algorithms demonstrated high data consistency (defined as >70% similarity) across all 44 species. SVM and RF achieved the highest levels of consistent transferences, successfully applying to 42 species in both hindcasting and forecasting scenarios. In contrast, GLM showed the lowest consistency, with consistent models for 35 species in hindcasting and 36 in forecasting (Table 1).

Niche models and model transferences. Binomial tests revealed that models calibrated with occurrences and climatic layers from the same period (auto1 and auto2) significantly deviated from random expectations for most species, indicating reliable model accuracy. Exceptions included GLM models for *Callospermophilus madrensis*, *Dasyprocta mexicana*, and *Dipodomys phillipsii* in T1 (1950-1979), and *Dipodomys phillipsii*, *Neotamias durangae*, *Neotoma goldmani*, and *Oryzomys guerrerensis* in T2 (1980-2009); a BRT model for *Pappogeomys bulleri* in T2, and a Maxent model for *Dasyprocta mexicana* in T2 (Supplementary Material, Table S4).

We found significant differences between algorithms for transferring models from T1 to T2 and vice versa ($H = 78.75$ and $H = 79.08$, respectively, both $p < 0.01$). RF consistently showed the highest mean TSS ($\bar{x} = 0.83 \pm 0.08$) in forecasting, while Maxent in hindcasting ($\bar{x} = 0.82 \pm 0.10$). Notably, Maxent showed significantly higher TSS values for hindcasting than forecasting ($w = 435$, $p = 0.002$). Conversely, BRT, RF, and Bioclim showed no significant directional differences. Bioclim recorded the lowest in both forecasting ($\bar{x} = 0.50 \pm 0.20$) and hindcasting ($\bar{x} = 0.52 \pm 0.21$).

Significant differences between algorithms were also evident in OI ($H = 27.27$, $p = 0.014$ in hinadcasting and $H = 15.93$, $p < 0.01$ in forecasting), with Bioclim ($\bar{x} = 0.52 \pm 0.21$) and GAM ($\bar{x} = 0.65 \pm 0.18$) showing the greatest variation in forecasting, and Maxent also displaying significant directional differences ($w = 382$, $p < 0.01$). RF maintained consistent performance across both directions (forecasting: $\bar{x} = 0.83 \pm 0.08$; hindcasting: $\bar{x} = 0.80 \pm 0.09$) (Fig. 2; Supplementary Material, Table S5).

The FPR varied significantly among algorithms in both forecasting and hindcasting (H = 92.237 and H = 77.102, both $p < 0.01$), with Bioclim showing the highest rates both in forecasting ($\bar{x} = 0.47 \pm 0.22$) and hindcasting ($\bar{x} = 0.47 \pm 0.22$), and RF the lowest (forecasting: $\bar{x} = 0.09 \pm 0.09$; hindcasting: $\bar{x} = 0.10 \pm 0.08$). Maxent showed significant differences in FPR between directions (w = 1062, p = 0.003). The FNR also showed significant differences between algorithms for hindcasting (H = 51.171, $p < 0.01$) and forecasting (H = 65.92, $p < 0.01$), where GAM recorded the highest values (Fig. 2; Supplementary Material, Table S5)..

Correlation analyses. Our correlation analyses between the number of occurrence records and TSS scores showed weak and non-significant positive relationships for most algorithms ($R^2 < 0.1$, $p > 0.05$). Exceptions were Bioclim, GAM, and GLM in hindcasting, with GLM also showing a marginal significance in forecasting (Fig. 3). In general, species with larger sample sizes generally achieved higher TSS scores,

Discussion

~~In this study, we explored the transferability of ecological niche models (ENMs) for Mexican endemic rodents using seven widely used algorithms across past and recent climatic conditions. Our findings reveal~~ that these algorithms differ in robustness when transferring models across time periods. Random Forest (RF), Maxent, Boosted Regression Trees (BRT), and Support Vector Machine (SVM) consistently performed well, while Bioclim showed the poorest performance. Notably, Maxent exhibited significant sensitivity to the direction of transfer, with marked differences between hindcasting and forecasting. Our results are consistent with previous studies that have observed variations in algorithm performance when transferring niche models to different climatic scenarios (Moreno-Amat et al., 2015; Beaumont et al., 2016; Bell & Schlaepfer, 2016; García-Callejas & Araújo, 2016; Liang et al., 2018; Qiao et al., 2019; Heikkinen, Marmion & Louto, 2012). However, to our knowledge, no prior studies have examined how occurrence data quality affects algorithm performance in both forecasting and hindcasting.

~~A well-known factor affecting model transferability~~ is the presence of non-analog climates between time periods (Sequeira et al., 2018; Essl et al., 2023). Climatic combinations outside the calibration scenario challenge all algorithms, particularly those with limited extrapolation capacity, such as Bioclim (Qiao et al., 2019). To assess the influence of non-analog climates on algorithm performance, we conducted a Mobility-Oriented Parity (MOP) analysis (Owens et al., 2013, which quantifies the multidimensional similarity between two climatic scenarios (calibration and transfer) and maps areas requiring strict extrapolation, with the *smop* package (Osorio-Olvera & Contreras Díaz, 2024) in R. Our results indicate that the areas with dissimilar climatic combinations—and thus where strict extrapolation is needed—are limited across Mexico: from T1 to T2, they occupy 0.58% of the country, and from T2 to T1, 1.7% (Supplementary Material, Fig. S1). Therefore, non-analog climates do not explain most of the observed variation in algorithm performance.

Another crucial factor affecting model performance is data quality, specifically the representation of the environmental combinations that define a species' ecological niche (van Proosdij et al., 2016; Jiménez-Valverde, 2020). Low-quality data may result from insufficient sampling or environmental bias (Wang & Jackson, 2023). We initially hypothesized that species with fewer occurrences would show poor transferability across all algorithms, as small sample sizes often lead to an incomplete ecological niche characterization. However, our results indicate that sample size impacts algorithms differently: while sample size minimally affected transferability for RF, BRT, and SVM, it significantly impacted Bioclim, GLM, GAM, and, to a lesser extent, Maxent. These findings align with previous studies where Maxent outperformed Bioclim (Hernandez et al., 2006) and GAM (Wang & Jackson, 2023) under small sample sizes.

Sample size also explains ~~the differing~~ performance ~~of~~ algorithms in forecasting and hindcasting. In our study, Maxent ~~exhibited~~ the greatest directional difference, showing greater robustness during hindcasting than forecasting (Fig. 2). Notably, Maxent's sensitivity to sample size was evident only in forecasting (Fig. 3), suggesting that small sample sizes more adversely affect its transferability than other algorithms like RF and BRT. This disparity in performance between Maxent and RF is consistent with Wang & Jackson's (2023) findings, who recommend RF for small sample sizes.

~~The robustness of modeling algorithms~~ ~~under~~ varying data qualities appears to hinge on their ability to accurately characterize ~~what García-Callejas & Araújo (2016) refer to as the geometrical complexity of the ecological niche~~—the structural characteristics of the boundary between suitable and unsuitable conditions in environmental space. For species with well-defined niches, like those specialized to specific environments, robust models are generally easier to obtain. In contrast, species with broader environmental preferences require more unbiased samples to accurately characterize their more complex niche boundaries. Consequently, some algorithms manage boundary complexity better than others, especially with small sample sizes. Our analysis revealed that Bioclim had the poorest performance in transferring niche models across temporal climatic scenarios. Bioclim, a simple environmental envelope model based on the range of values from occurrence records in the predictor variables (Nix, 1986), is highly sensitive to extreme values and the number of predictors (Beaumont, Hughes & Poulsen, 2005). Additionally, Bioclim’s quadrangular representation of the ecological niche in environmental space limits its ability to model complex niche geometries. A similar issue may arise with GLM, particularly with small sample sizes (Guisan, Edwards & Hastie, 2002).

In contrast, RF and Maxent exhibited the ~~highest~~ transfer capacity, followed closely by BRT and SVM. RF also ~~demonstrated~~ the most consistent performance between transfer directions. This algorithm has proven robust for transferring models with both virtual (García-Callejas & Araújo, 2016) and real (Mi et al., 2017) species, ~~shows~~ good interpolation performance (Bell & Schlaepfer, 2016; Liang et al., 2018), and relatively ~~low~~ overprediction (Mi et al., 2017). However, this robustness comes at the cost of overfitting, which sometimes limits its extrapolation capacity beyond the calibration range (Heikkinen, Marmion & Lousteau, 2012). Maxent, on the other hand, is less prone to overfitting, particularly when parameterized *ad hoc* for specific species (Merow, Smith & Silander, 2013). In summary, the four machine-learning algorithms generally outperformed the two regression-based and climatic envelope algorithms.

However, as highlighted in numerous comparative studies, there is no "silver bullet" algorithm that consistently performs best across all data structures (Qiao, Soberón & Peterson, 2015). This is even more evident in model transferability, where algorithm weaknesses are amplified (Pearson et al., 2006; Moreno-Amat et al., 2015).

A final note of caution is that while our analyses were designed to evaluate algorithm performance for temporal transferability using Mexican endemic rodents as the model system, a potential source of error lies in the continuously updated taxonomy of these species. Our analyses were based on the latest revision of Mexican mammals (Ramírez-Pulido et al., 2014); however, recent proposals suggest species separation for *Peromyscus melanophrys*, *P. furvus*, *P. levipes*, *P. zarhynchus*, and *Osgoogomys banderanus* (Lorenzo et al., 2016; Almendra et al., 2018; Cruz-Gómez et al., 2021; Bradley et al., 2022). These taxonomic changes could affect the quantity and spatial structure of occurrences for these species, potentially impacting model performance (Soley-Guardia, Serrano & Anderson, 2024).

Conclusions

Ecological niche modeling is often used to transfer models across temporal scenarios for climate change analysis. However, the suitability of the species for such transfers and the robustness of the chosen algorithms are often overlooked. Our results highlight~~s~~ that the performance of these algorithms, and consequently the reliability of temporal transfers, is primarily influenced by the quality of data. ~~Low~~ sample sizes ~~significantly compromise~~ the effectiveness of model transfers, with some algorithms being more affected than others. Indeed, an algorithm may yield different results for the same species when transferring models to past versus future scenarios, indicating that not all species and algorithms are equally suited for transferring models across temporal scenarios. Among the algorithms evaluated, those capable of modeling complex ecological boundaries with minimal overfitting—such as Random Forest, Maxent, and Boosted Regression Trees—consistently outperformed simpler algorithms ~~like~~

Bioclim ~~or~~ GLM. Consequently, we strongly recommend a careful assessment of both species and algorithms before proceeding with temporal transfers. In this regard, the retrospective cross-temporal approach presented here offers a valuable alternative.

Acknowledgements

This work is fulfillment of CNMA's Graduate Doctoral Degree program in Biological Sciences at the Universidad Nacional Autónoma de México (UNAM). CNMA received logistical support from the Posgrado en Ciencias Biológicas and Instituto de Biología-UNAM. We express our gratitude to Livia León-Paniagua and Fausto Mendez for their guidance during the study, to Luis Osorio-Olvera for his assistance with the MOP analysis, Town Peterson for his review to an earlier version of the manuscript, and James Roper, Iván Ray-Rodríguez and two anonymous reviewers for their valuable comments on the manuscript.

References

Allouche, O., Tsoar, A. & Kadmon, R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43:1223–1232. DOI: 10.1111/j.1365-2664.2006.01214.x.

Almendra, A.L., González-Cózatl, F.X., Engstrom, M.D., Rogers, D.S. 2018. Evolutionary relationships and climatic niche evolution in the genus *Handleyomys* (Sigmodontinae: Oryzomyini). *Molecular Phylogenetics and Evolution* 128:12–25. DOI: 10.1016/j.ympev.2018.06.018.

Anderson, R.P., Lew, D., Peterson, A.T. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* 162:211–232. DOI: 10.1016/S0304-3800(02)00349-6.

Araújo, M.B. & Peterson, A.T. 2012. Uses and misuses of bioclimatic envelope modeling. *Ecology* 93:1527–1539.

Araújo, M., Thuiller, W. & Pearson, R.G. 2006. Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography* 33:1712–1728. DOI: 10.1111/j.1365-2699.2006.01482.x.

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

Barbet-Massin, M. & Jetz, W. 2014. A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Diversity and Distributions*:1–11. DOI: 10.1111/ddi.12229.

Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* 3:327–338. DOI: 10.1111/j.2041-210X.2011.00172.x.

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 modeling and species distribution modeling. *Ecological Modelling* 222:1810–1819. DOI: 10.1016/j.ecolmodel.2011.02.011.

Beaumont, L.J., Graham, E., Duursma, D.E., Wilson, P.D., Cabrelli, A., Baumgartner, J.B., Hallgren, W., Esperón-Rodríguez, M., Nipperess, D.A., Warren, D.L., Laffan, S.W., VanDerWal, J. 2016. Which species distribution models are more (or less) likely to project broad-scale, climate-induced shifts in species ranges? *Ecological Modelling* 342:135–146. DOI: https://doi.org/10.1016/j.ecolmodel.2016.10.004

Beaumont, L.J., Hughes, L. & Poulsen, M. 2005. Predicting species distributions: Use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling* 186:250–269. DOI: 10.1016/j.ecolmodel.2005.01.030.

Bell, D.M. & Schlaepfer, D.R. 2016. On the dangers of model complexity without ecological justification in species distribution modeling. *Ecological Modelling* 330:50–59. DOI: 10.1016/j.ecolmodel.2016.03.012.

Booth, T.H., Nix, H.A., Busby, J.R., Hutchinson, M.F. 2014. Bioclim: The first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. *Diversity and Distributions* 20:1–9. DOI: 10.1111/ddi.12144.

Bradley, R.D., Ordóñez-Garza, N., Thompson, C.W., Wright, E.A., Ceballos, G., Kilpatrick, C.W., Schmidly, D.J. 2022. Two new species of *Peromyscus* (Cricetidae: Neotominae) from the Transverse Volcanic Belt of Mexico. *Journal of Mammalogy* 103:255–274. DOI: 10.1093/jmammal/gyab128.

Breiman L. 2001. Random forests. *Machine Learning* 45:5–32. DOI: 10.1007/978-3-662-56776-0_10.

Brotons, L., Thuiller, W., Araújo, M.B., Hirzel, A.H. 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27:437–448. DOI: https://doi.org/10.1111/j.0906-7590.2004.03764.x

Charney, N.D., Record, S., Gerstner, B.E., Merow, C., Zarnetske, P.L., Enquist, B.J. 2021. A test of species distribution model transferability across environmental and geographic space for 108 western North American tree species. *Frontiers in Ecology and Evolution*, *9*, 689295. DOI: https://doi.org/10.3389/fevo.2021.689295

Cruz-Gómez, A., Castro-Campillo, A., Ávila-Valle, Z.A., León-Paniagua L., Ramírez-Sánchez, M., Ramírez-Pulido, J. 2021. Rejection of the monotypic status of *Peromyscus furvus* (Rodentia: Cricetidae), with consequences for its species group. *Therya* 12:347–367. DOI: 10.12933/therya-21-1122.

Cuervo-Robayo, Á., Ureta, C., Gómez-Albores, M.A., Meneses Mosquera, A.K., Téllez Valdés, O., Martínez-Meyer, E. 2020. One hundred years of climate change in Mexico. *PLoS ONE* 15:1–19. DOI: 10.1074/jbc.RA118.003614.

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F.T., D'Amen, M., Randin, C., Engler, R., Pottier, J., Pio, D., Dubuis, A., Pellissier, L., Mateo, R.G., Hordijk, W., Salamin, N., Guisan, A. 2017. Ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography* 40:774–787. DOI: 10.1111/ecog.02671.

Dobrowski, S.Z, Thorne, J.H, Greenberg, J.A, Safford, H.D, Mynsberge, A.R, Crimmins, S.M, Swanson, A.K. 2011. Modeling plant ranges over 75 years of climate change in California, USA: Temporal transferability and species traits. *Ecological Monographs* 81:241–257. DOI: 10.1890/10-1325.1.

Drake, J.M., Randin, C. & Guisan, A. 2006. Modelling ecological niches with support vector machines. *Journal of Applied Ecology* 43:424–432. DOI: 10.1111/j.1365-2664.2006.01141.x.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M.C., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., Zimmermann, N.E. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151. DOI: https://doi.org/10.1111/j.2006.0906-7590.04596.x

Elith, J., Leathwick, J.R. & Hastie, T. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology* 77:802–813. DOI: 10.1111/j.1365-2656.2008.01390.x.

Essl, F., García-Rodríguez, A., Lenzner, B.M.A., Capinha, C., Gaüzère, P., Guisan, A., Kühn, I., Lenoir, J., Richardson, D.M., Rumpf, S.B., Svenning, J-C., Thuiller, W., Zurell, D., Dullinger, S. 2023. Potential sources of time lags in calibrating species distribution models. *Journal of Biogeography*:1–14. DOI: 10.1111/jbi.14726.

Fan, J.Y., Zhao, N.X., Li, M., Gao, W.F., Wang, M.L., Zhu, G.P. 2018. What are the best predictors for invasive potential of weeds? Transferability evaluations of model predictions based on diverse environmental data sets for Flaveria bidentis. *Weed Research* 58:141–149. DOI: 10.1111/wre.12292.

Fielding, A.H. & Bell, J.F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24:38–49. DOI: 10.1017/S0376892997000088.

García-Callejas, D. & Araújo, M.B. 2016. The effects of model and data complexity on predictions from species distributions models. *Ecological Modelling* 326:4–12. DOI: 10.1016/j.ecolmodel.2015.06.002.

Garcia, R.A., Cabeza, M., Altwegg, R., Araújo, M.B. 2016. Do projections from bioclimatic envelope models and climate change metrics match? *Global Ecology and Biogeography* 25:65–74. DOI: 10.1111/geb.12386.

Golicher, D., Ford, A., Cayuela, L., Newton, A. 2012. Pseudo-absences, pseudo-models and pseudo-niches: pitfalls of model selection based on the area under the curve. *International Journal of Geographical Information Science* 26:2049–2063. DOI: 10.1080/13658816.2012.719626.

Guisan, A., Edwards, T.C. & Hastie, T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* 157:89–100. DOI: 10.1016/S0304-3800(02)00204-1.

Habibullah, M.S., Din, B.H., Tan, S.H., Zahid, H. 2022. Impact of climate change on biodiversity loss: global evidence. *Environmental Science and Pollution Research* 29:1073–1086. DOI: 10.1007/s11356-021-15702-8.

Heikkinen, R.K., Marmion, M. & Luoto, M. 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography* 35:276–288. DOI: 10.1111/j.1600-0587.2011.06999.x.

Hernandez, P.A., Graham, C.H., Master, L.L., Albert, D.L. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29:773–785. DOI: 10.1111/j.0906-7590.2006.04700.x.

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. DOI: 10.1002/joc.1276.

Hijmans, R.J. & Graham, C. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology* 12:2272–2281. DOI: 10.1111/j.1365-2486.2006.01256.x.

Hijmans, A.R.J., Phillips, S., Leathwick, J., Elith, J., Hijmans, M.R.J. 2017. Package ‘DISMO’ *CRAN*. Available at https://cran.r-project.org/web/packages/dismo/index.html (accessed 18 may 2022).

Jaroszynska, F., Rixen, C., Woodin, S., Lenoir, J., Wipf, S. 2023. Resampling alpine herbarium records reveals changes in plant traits over space and time. *Journal of Ecology* 111:338–355. DOI: 10.1111/1365-2745.14062.

Jiménez-Valverde, A. (2020). Sample size for the evaluation of presence-absence models. *Ecological Indicators*, *114*, 106289. https://doi.org/10.1016/j.ecolind.2020.106289

Kass, J.M., Muscarella, R., Galante, P.J., Bohl, C.L., Pinilla‐Buitrago, G.E., Boria, R.A., Solely-Guardia, M., Anderson, R.P. (2021). ENMeval 2.0: Redesigned for customizable and reproducible modeling of species’ niches and distributions. *Methods in Ecology and Evolution*, *12*(9), 1602-1608.

Kharouba, H.M., Algar, A.C. & Kerr, J.T. 2009. Historically calibrated predictions of butterfly species’ range shift using global change as a pseudo-experiment. *Ecology* 90:2213–2222. DOI: 10.1890/08-1304.1.

Liang, W., Papeş, M., Tran, L., Grant, J., Washington-Allen, R., Stewart, S., Wiggins, G. 2018. The effect of pseudo-absence selection method on transferability of species distribution models in the context of non-adaptive niche shift. *Ecological Modelling* 388:1–9. DOI: 10.1016/j.ecolmodel.2018.09.018.

Lorenzo, C., Álvarez-Castañeda, S.T., Pérez-Consuegra, S.G., Patton, J.L. 2016. Revision of the Chiapan deer mouse, *Peromyscus zarhynchus*, with the description of a new species. *Journal of Mammalogy* 97:910–918. DOI: 10.1093/jmammal/gyw018.

MacLean, S.A., Rios Dominguez, A.F., de Valpine, P., Beissinger, S.R. 2018. A century of climate and land-use change cause species turnover without loss of beta diversity in California’s Central Valley. *Global Change Biology* 24:5882–5894. DOI: 10.1111/gcb.14458.

Maiorano, L., Falcucci, A., Zimmermann, N.E., Psomas, A., Pottier, J., Baisero, D., Rondinini, C., Guisan, A., Boitani, L. 2011. The future of terrestrial mammals in the Mediterranean basin under climate change. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366:2681–2692. DOI: 10.1098/rstb.2011.0121.

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.

Merow, C., Smith, M.J., Edwards, Jr TC., Guisan, A., McMahon, S.M., Normand, S., Thuiller, W., Wüest, R.O., Zimmermann, N.E., Elith, J. 2014. What do we gain from simplicity versus complexity in species distribution models? *Ecography* 37:1–15. DOI: 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. DOI: 10.1111/j.1600-0587.2013.07872.x.

Mi, C., Huettmann, F., Guo, Y., Han, X., Wen, L. 2017. Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ*. DOI: 10.7717/peerj.2849.

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. DOI: 10.1016/j.ecolmodel.2015.05.035.

Naimi, B. & Araújo, M.B. 2016. SDM: A reproducible and extensible R platform for species distribution modelling. *Ecography* 39:368–375. DOI: 10.1111/ecog.01881.

Nix, H. A., & Busby, J. 1986. BIOCLIM, a bioclimatic analysis and prediction system. *Division of Water and Land Resources: Canberra*.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience* 51:933. DOI: 10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

Osorio-Olvera, L & Contreras-Díaz, R.G. 2024. smop. Simple but Fast MOP Analysis ver. 0.0.2. https://luismurao.github.io/smop/.

Owens, H. L., Campbell, L. P., Dornak, L. L., Saupe, E. E., Barve, N., Soberón, J.,Ingenloff K., Lira-Noriega A., Hensz C.M., Myers C.E.,Peterson AT. 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.

Pacifici, M., Foden, W.B., Visconti, P., Watson, J.E.M., Butchart, S.H.M., Kovacs, K.M., Scheffers, B.R., Hole, D.G., Martin, T.G., Akçakaya, H.R., Corlett, R.T., Huntley, B., Bickford, D., Carr, J.A., Hoffmann, A.A., Midgley, G.F., Pearce-Kelly, P., Pearson, R.G., Williams, S.E., Willis, S.G., Young, B., Rondinini, C. 2015. Assessing species vulnerability to climate change. *Nature Climate Change* 5:215–225. DOI: 10.1038/nclimate2448.

Parmesan, C. & Yohe, G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421:37–42. DOI: 10.1038/nature01286.

Pearson, R.G. 2006. Climate change and the migration capacity of species. *Trends in Ecology and Evolution* 21:111–113. DOI: 10.1016/j.tree.2005.11.022.

Pearson, R.G. & Dawson, TP. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* 12:361–371. DOI: 10.1046/j.1466-822X.2003.00042.x.

Peterson, A.T. 2011. Ecological niche conservatism: A time-structured review of evidence. *Journal of Biogeography* 38:817–827. DOI: 10.1111/j.1365-2699.2010.02456.x.

Peterson, A.T., Martínez-Meyer, E. & González-Salazar, C. 2004. Reconstructing the Pleistocene geography of the Aphelocoma jays (Corvidae). *Diversity and Distributions* 10:237–246. DOI: https://doi.org/10.1111/j.1366-9516.2004.00097.x.

Peterson, A.T., Navarro-Siguenza, A.G., Martínez-Meyer, E., Cuervo-Robayo, A.P., Berlanga, H., Soberón, J. 2015. Twentieth century turnover of Mexican endemic avifaunas: Landscape change versus climate drivers. *Science Advances* 1:e1400071–e1400071. DOI: 10.1126/sciadv.1400071.

Phillips, S.J. 2008. Transferability, sample selection bias and background data in presence-only modelling: a response to Peterson et al. (2007). *Ecography*:080227084236895. DOI: 10.1111/j.2007.0906-7590.05378.x.

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

Piirainen, S., Lehikoinen, A., Husby, M., Kålås, J.A., Lindström, Å., Ovaskainen, O. 2023. Species distributions models may predict accurately future distributions but poorly how distributions change: A critical perspective on model validation. *Diversity and Distributions* 29:654–665. DOI: 10.1111/ddi.13687.

Pohlert, T. 2016. The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR). *CRAN*:1–27.

Prasad, A.M., Iverson, L.R. & Liaw, A. 2006. Newer Classification and Regression Tree Techniques : Bagging and Random Forests for Ecological Prediction. *Ecosystems* 9:181–199. DOI: 10.1007/s10021-005-0054-1.

Qiao, H., Soberón, J. & Peterson, A.T. 2015. No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. *Methods in Ecology and Evolution*, *6*(10), 1126-1136.

Qiao, H., Feng, X., Escobar, L.E., Peterson, A.T., Soberón, J., Zhu, G., Papeş, M. 2019. An evaluation of transferability of ecological niche models. *Ecography* 42:521–534. DOI: 10.1111/ecog.03986.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. DOI: 10.1007/978-3-540-74686-7.

Radosavljevic, A. & Anderson, R.P. 2014. Making better Maxent models of species distributions: Complexity, overfitting and evaluation. *Journal of Biogeography* 41:629–643. DOI: 10.1111/jbi.12227.

Ramírez-Pulido, J., González-Ruiz, N., Gardner, A.L., Arroyo-Cabrales, J. 2014. *List of recent land mammals of Mexcio*. Lubbock, Texas. DOI: 10.1177/002194368302000404.

Rebelo, H., Tarroso, P. & Jones, G. 2010. Predicted impact of climate change on European bats in relation to their biogeographic patterns. *Global Change Biology* 16:561–576. DOI: 10.1111/j.1365-2486.2009.02021.x.

Regos, A., Gagne L., Alcaraz-Segura, D., Honrado, J.P., Domínguez, J. 2019. Effects of species traits and environmental predictors on performance and transferability of ecological niche models. *Scientific Reports* 9:1–14. DOI: 10.1038/s41598-019-40766-5.

Rubidge, E.M., Monahan, W.B., Parra, J.L., Cameron, S.E., Brashares, J.S. 2011. The role of climate, habitat, and species co-occurrence as drivers of change in small mammal distributions over the past century. *Global Change Biology* 17:696–708. DOI: 10.1111/j.1365-2486.2010.02297.x.

Sequeira, A.M.M., Bouchet, P.J., Yates, K.L., Mengersen, K., Caley, M.J. 2018. Transferring biodiversity models for conservation: Opportunities and challenges. *Methods in Ecology and Evolution* 9:1250–1264. DOI: 10.1111/2041-210X.12998.

Sillero, N., Arenas-Castro, S., Enriquez-Urzelai, U., Vale, C. G., Sousa-Guedes, D., Martínez-Freiría, F., Real, R., Barbosa, A. M. 2021. Want to model a species niche? A step-by-step guideline on correlative ecological niche modelling. *Ecological Modelling*, *456*, 109671. DOI: https://doi.org/10.1016/j.ecolmodel.2021.109671

Sillero, N., Campos, J.C., Arenas-Castro, S., Barbosa, A.M. 2023. A curated list of R packages for ecological niche modelling. *Ecological Modelling* 476. DOI: 10.1016/j.ecolmodel.2022.110242.

Soberón, J. & Nakamura, M. 2009. Niches and distributional areas : Concepts, methods and assumptions. *PNAS* 106:19644–19650.

Soley‐Guardia, M., Alvarado‐Serrano, D.F. & Anderson, R.P. 2024. Top ten hazards to avoid when modeling species distributions: a didactic guide of assumptions, problems, and recommendations. *Ecography*, *2024*(4), e06852. DOI: 10.1111/ecog.06852

The Document Foundation. (2020). LibreOffice 7.0 (Version 7) [Java, Python, C++; Linux]. The Document Foundation. Download LibreOffice | LibreOffice - Free Office Suite - Based on OpenOffice - Compatible with Microsoft.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., De Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., Van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L.,

van Proosdij, A. S., Sosef, M. S., Wieringa, J. J., Raes, N. (2016). Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, *39*(6), 542-552.https://doi.org/10.1111/ecog.01509

Vaz, U.L., Cunha, H.F. & Nabout, J.C. 2015. Trends and biases in global scientific literature about ecological niche models. *Brazilian Journal of Biology* 75:17–24. DOI: 10.1590/1519-6984.22713.

Waltari, E. & Guralnick, R.P. 2009. Ecological niche modelling of montane mammals in the Great Basin, North America: examining past and present connectivity of species across basins and ranges. *Journal of Biogeography* 36:148–161. DOI: 10.1111/j.1365-2699.2008.01959.x.

Wang, L., & Jackson, D. A. (2023). Effects of sample size, data quality, and species response in environmental space on modeling species distributions. *Landscape Ecology*, *38*(12), 4009-4031. https://doi.org/10.1007/s10980-023-01771-2

Warren, D.L., Wright, A.N., Seifert, S.N., Shaffer, H.B. 2014. Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90

California vertebrate species of concern. *Diversity and Distributions* 20:334–343. DOI: 10.1111/ddi.12160.

Warton, D. & Aarts, G. 2013. Advancing our thinking in presence-only and used-available analysis. *Journal of Animal Ecology* 82:1125–1134. DOI: 10.1111/1365-2656.12071.

Werkowska, W., Márquez, AL., Real, R., Acevedo, P. 2017. A practical overview of transferability in species distribution modeling. *Environmental Reviews* 25:127–133. DOI: 10.1139/er-2016-0045.

Widick, I.V. & Bean, W.T. 2019. Evaluating current and future range limits of an endangered, keystone rodent (*Dipodomys ingens*). *Diversity and Distributions* 25:1074–1087. DOI: 10.1111/ddi.12914.

Wiens, J.A., Stralberg, D., Jongsomjit, D., Howell, C.A., Snyder, M.A. 2009. Niches, models, and climate change: assessing the assumptions and uncertainties. *PNAS* 106 Suppl:19729–19736. DOI: 10.1073/pnas.0901639106.

Yates, K.L., Bouchet, P.J., Caley, M.J., Mengersen, K., Randin, C.F., Parnell, S., Fielding, A.H., Bamford, A.J., Ban, S., Barbosa, A.M., Dormann, C.F., Elith, J., Embling, C.B., Ervin, G.N., Fisher, R., Gould, S., Graf, R.F., Gregr, E.J., Halpin, P.N., Heikkinen, R.K., Heinänen, S., Jones, A.R., Krishnakumar, P.K., Lauria, V., Lozano-Montes, H., Mannocci, L., Mellin C., Mesgaran, M.B., Moreno-Amat, E., Mormede, S., Novaczek, E., Oppel, S., Ortuño Crespo, G., Peterson, A.T., Rapacciuolo, G., Roberts, J.J., Ross, R.E., Scales, K.L., Schoeman, D., Snelgrove, P., Sundblad, G., Thuiller, W., Torres, L.G., Verbruggen, H., Wang, L., Wenger, S., Whittingham, M.J., Zharikov, Y., Zurell, D., Sequeira, A.M.M. 2018. Outstanding Challenges in the Transferability of Ecological Models. *Trends in Ecology and Evolution* 33:790–802. DOI: 10.1016/j.tree.2018.08.001.

Zhu, G., Fan, J., & Peterson, A. T. 2021. Cautions in weighting individual ecological niche models in ensemble forecasting. *Ecological Modelling*, *448*, 109502.DOI: https://doi.org/10.1016/j.ecolmodel.2021.109502.

Table 1(on next page)

Species with sufficient occurrences

Number of species with consistent occurrence data between time periods for each modeling algorithm

Table 1. Number of species with consistent occurrence data suitable for temporal transferences by modeling algorithm.

Algorithm	**Hindcasting**	**Forecasting**
Bioclim	40	38
Boosted Regression Trees (BRT)	39	39
Generalized Additive Models (GAM)	41	41
Generalized Linear Models (GLM)	35	36
Maxent	38	38
Random Forest (RF)	42	42
Support Vector Machine (SVM)	42	42

Figure 1

Methodological approach to evaluate algorithm performance over time.

Figure 1. Methodological approach applied to evaluate the ability of algorithms to transfer niche models over time. (A) Ecological niche model calibrated with occurrence data and environmental variables from the same period. (B) Ecological niche model calibrated with occurrence data from one period and environmental variables from a different period. (C) Model transferred to a different period. (D) Geographic validation by comparing the model transferred from one period against the model calibrated in the other period. The example maps correspond to the cotton rat *Sigmodon mascotensis*, and the climate data for the analysis were obtained from Cuervo-Robayo et al. (2020). The figure was created with Libreoffice-Impress 7.1.1.2 (The Document Foundation, 2020).

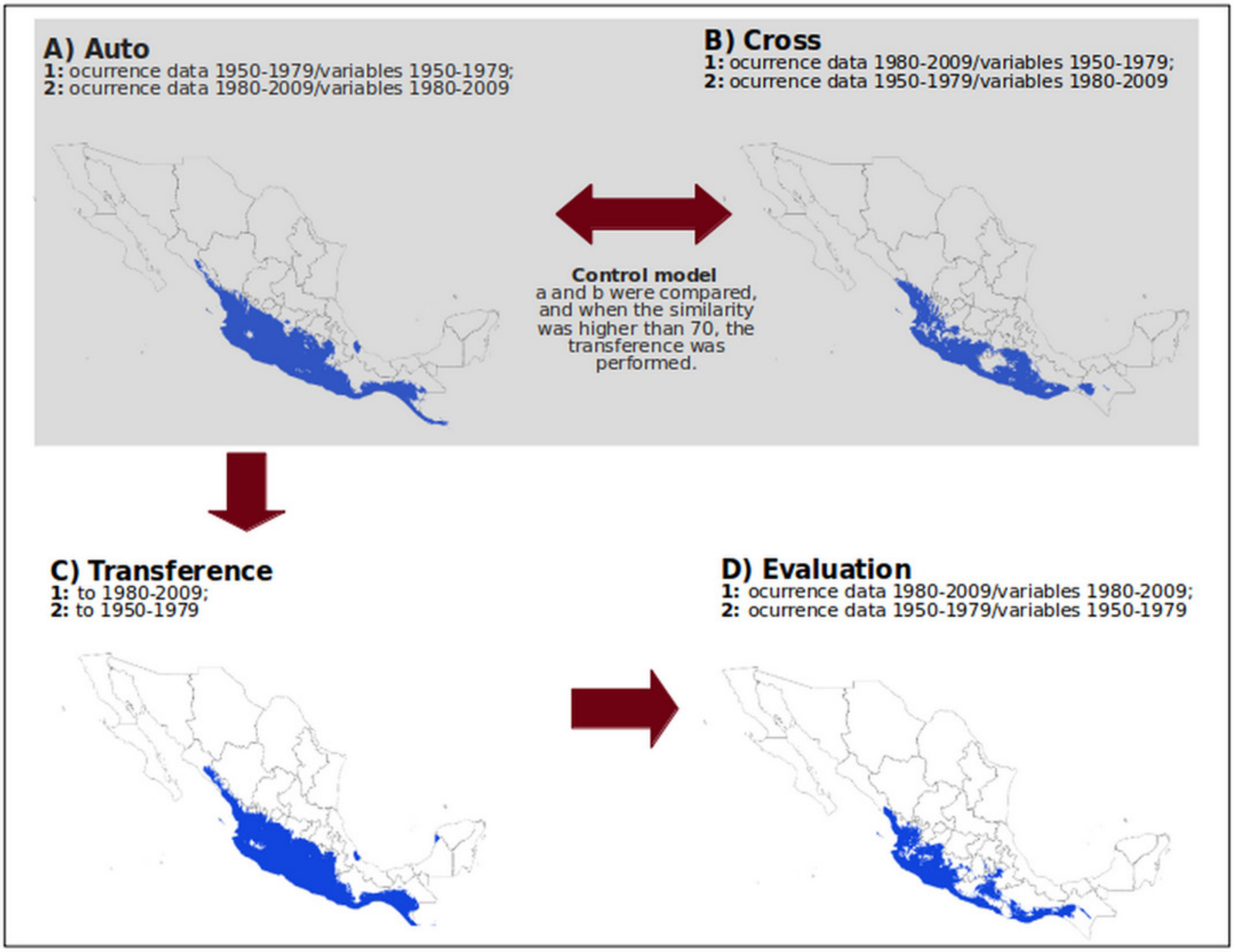

A) Auto
1: ocurrence data 1950-1979/variables 1950-1979;
2: ocurrence data 1980-2009/variables 1980-2009
B) Cross
1: ocurrence data 1980-2009/variables 1950-1979;
2: ocurrence data 1950-1979/variables 1980-2009
Control model
a and b were compared, and when the similarity was higher than 70, the transference was performed.
C) Transference
1: to 1980-2009;
2: to 1950-1979
D) Evaluation
1: ocurrence data 1980-2009/variables 1980-2009;
2: ocurrence data 1950-1979/variables 1950-1979

Figure 2

Algorithm performance measured with different metrics.

Figure 2. Performance of niche modeling algorithms for hindcasting and forecasting using different metrics: True Skill Statistics, TSS; Overlap Index OI; False Positive Rate, FPR; and False Negative Rate, FNR. Solid red and blue dots represent the median and the upper and lower bars, the interquartile range, and the width of each plot represents the density of observations.

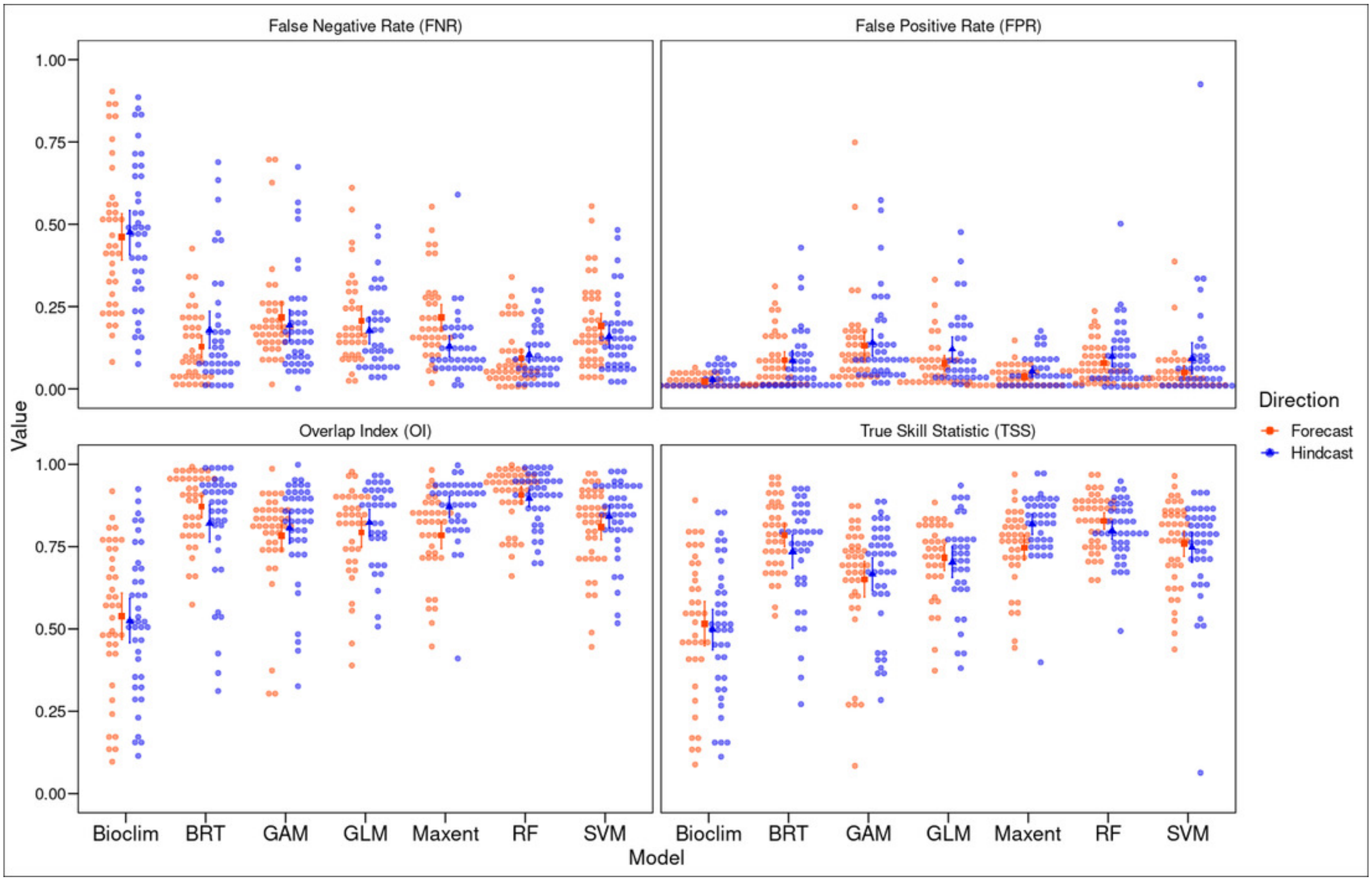
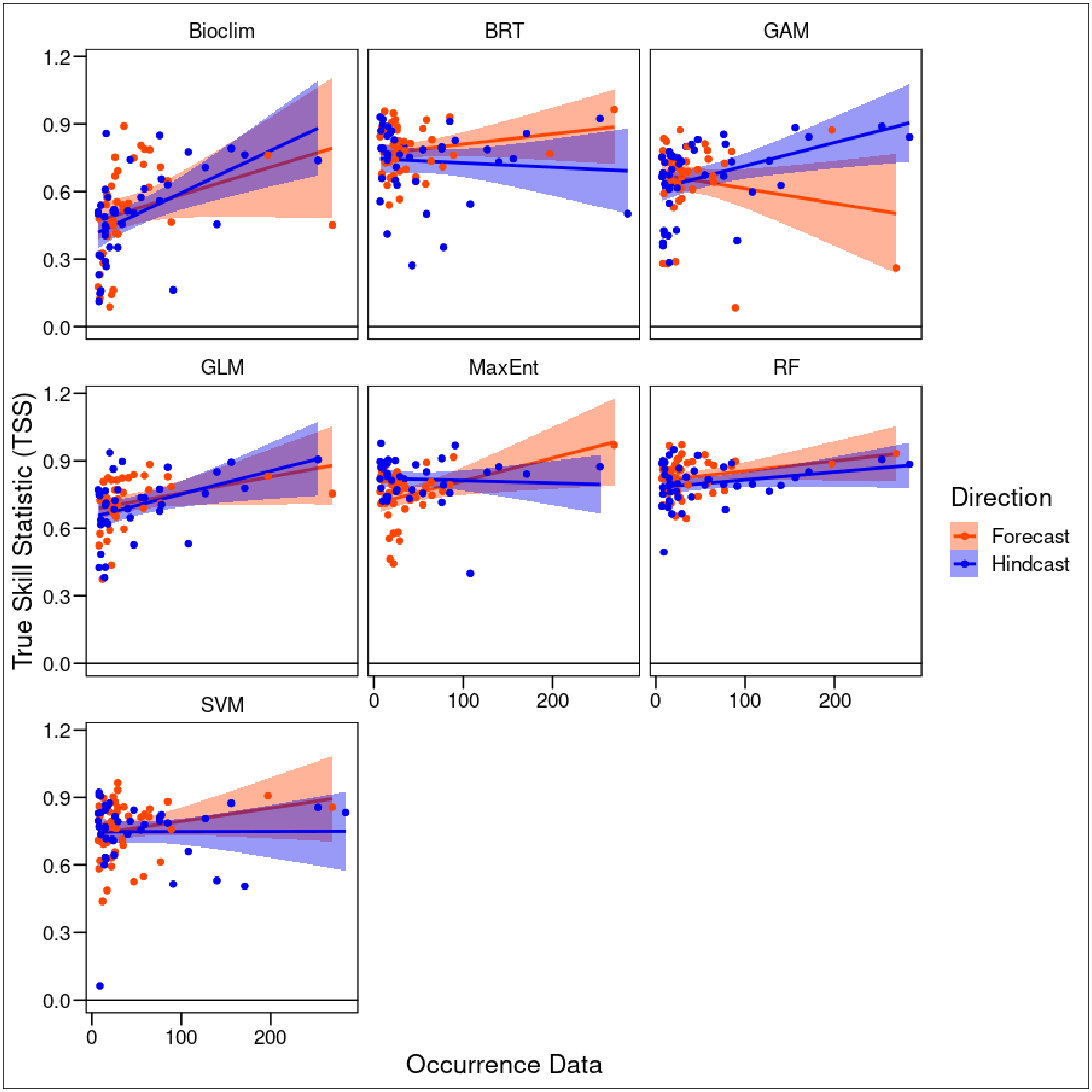


Figure 3

Correlation between the number of occurrence records and model performance by True Skill Statistics (TSS).

Figure 3. Pearson correlation analyses between the number of occurrence records of each species and model performance measured by True Skill Statistics (TSS) for each algorithm and direction. Dots represent individual species and lines the linear trend, red indicates forecasting and blue hindcasting.


Bioclim
BRT
GAM
GLM
MaxEnt
RF
SVM
1.2
0.9
0.6
0.3
0.0
0
100
200
True Skill Statistic (TSS)
Occurrence Data
Direction
Forecast
Hindcast