A retrospective approach for evaluating transference of ecological niche modeling ~~evaluating ecological niche modeling transferences~~ over time: The case of Mexican endemic rodents

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

[1] Posgrado de Ciencias Biológicas, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico.
[2] Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico.
~~[3] Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico.~~

Comentado [IR1]: Is it the same as number 2, it should be different?

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

Abstract

Ecological niche modeling (ENM) is an approach to ~~infer~~ predict the suitable conditions for species' persistence and their potential geographic distributions. ENM is extensively used to assess the potential effects of climate change on species' distributions, although modeling algorithms are recognized ~~acknowledged~~ as an important source of uncertainty in climate change projections. A problem is that these models cannot be properly assessed in the context of future projections. 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)) for transferring ENM over time for Mexican endemic rodents. To accomplish this, we employed~~To do so, we followed~~ a retrospective approach by transferring models from the near past (1950-1979) to the present (1980-2009) and vice versa. Our findings yielded three key ~~We~~

Comentado [IR2]: Is it a problem, maybe use a sentence like: The main challenge….

~~found three important~~ results: (1) The quality of input data and the algorithm had a significant effect on model output and performance; (2) algorithm performance was different for transferring models to the future than to the past; and (3) the most robust algorithms were RF, BRT and Maxent, whereas Bioclim was the least consistent. In conclusion, algorithm choice is critical for transferring ENM over time. Since no algorithm performed consistently better than the rest, we recommend ~~to~~ testing different algorithms *prior* transferring models to future scenarios under a retrospective approach.

Comentado [IR3]: Choice of algorithm?

Comentado [IR4]: More results and conclusions

Introduction

Current climate change is affecting biodiversity in a myriad of ways (Parmesan & Yohe, 2003; Bellard et al., 2012; MacLean et al., 2018; Widick & Bean, 2019; Cuervo-Robayo et al., 2020; Habibullah et al., 2022) and biotas are changing due to alterations in the abundance and distributional shifts of species (Moritz et al., 2008; Kharouba, Algar & Kerr, 2009; Dobrowski et al., 2011; Pacifici et al., 2015; Peterson et al., 2015b; Jaroszynska et al., 2023). For some years now, ecological niche modeling (ENM) is a popular tool for analyzing geographic distributions of species in the context of past and current climatic changes (Araújo et al., 2005; Beaumont et al., 2009; García-Callejas & Araújo, 2016; Sequeira et al., 2018; Yates et al., 2018; García ~~-~~et al., 2023). ENM includes a suite of concepts and methodologies based on ecological niche theory that has taken advantage of other disciplines, such as informatics, geographic information systems (GIS), and statistics (Austin, 2007; Jiménez-Valverde, Lobo & Hortal, 2008; Soberón & Nakamura, 2009; Wiens et al., 2009).

Comentado [IR5]: Not in the reference list, please check it.

ENMs are correlative methods that require georeferenced species' occurrence data and a set of predictor variables in the form of GIS raster layers to reconstruct the ecological characteristics that define the conditions where species inhabit (i.e., its ecological niche) and then project these conditions onto the landscape to produce a geographical expression of the species' niche, commonly referred to as its potential distribution (Elith et al., 2006; Peterson, 2011; Sillero, 2011). Several methods have been implemented to produce niche models, which can be classified according to the type of input occurrence data as 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; Warton & Aarts, 2013; Fan et al., 2018; Qiao et al., 2019; Sillero et al., 2023). From these, presence-absence methods are generally more robust when unequivocal absence data exist (Brotons et al., 2004; Golicher et al., 2012; Merow et al., 2014); however, in most cases, reliable absences are lacking. Therefore, presence-only, presence-pseudoabsence and, presence-background algorithms have become popular in the last two decades (Vaz, Cunha & Nabout, 2015).

Comentado [IR6]: Because.. could you explain why?

ENM has been broadly used to evaluate the potential geographic consequences of climatic changes by projecting the modeled species' niche under current climatic conditions to past or future climatic scenarios, a process also known as model transference (Pearson & Dawson, 2003; Peterson, Martínez-Meyer & González-Salazar, 2004; Thomas et al., 2004; Waltari & Guralnick, 2009; Heikkinen, Marmion & Luoto, 2012; Garcia et al., 2016). At least three main assumptions are behind model transferences over time that may or may not hold true: (1) the ecological niche does not evolve during the timeframe when the transference is performed (Soberón & Nakamura, 2009); (2) during a climatic change, even if the interactions among variables change, the relationship of those variables with the species remains constant (Hijmans & Graham, 2006; Wiens et al., 2009); and (3) the species is in equilibrium with the environment in the calibration time/area, meaning that they occupy the available suitable areas that are accessible to them (Sequeira et al., 2018; Yates et al., 2018).

Comentado [IR7]: Please, rewrite this sentence

Furthermore, a key issue in model transference is the effect of non-analog climates (i.e., the set of environmental conditions that exist in an alternative climatic scenario different from the calibration one) on model performance (Sequeira et al., 2018). Non-analog climates impose a challenge to algorithms because no calibration exists for those climates, thus, when algorithms face non-analogues in alternative climatic scenarios (past or future), algorithms respond differently according to the instructions of the program, producing contrasting or even contradicting results (Pearson, 2006; Araújo & Peterson, 2012; Essl et al., 2023). This is particularly problematic in future projections because there are no occurrence data that can be used to evaluate algorithm performance. Thus, one solution to evaluate algorithm performance in climate change studies is to project models between two time periods for which they have occurrence data, for example, from the relatively recent past to the present (Rubidge et al., 2011; Piirainen et al., 2023).

Several studies have focused on evaluating the reliability of niche models for a specific time (Beaumont & Hughes, 2002; Araújo, Thuiller & Pearson, 2006; Maiorano et al., 2011; Werkowska et al., 2017; Sequeira et al., 2018), but only a few have analyzed the performance of algorithms for transferring models over time (Prasad, Iverson & Liaw, 2006; Hijmans & Graham, 2006; Kharouba, Algar & Kerr, 2009; Rubidge et al., 2011; Moreno-Amat et al., 2015; García-Callejas & Araújo, 2016), with performance results spanning from good to bad for different species and algorithms, making it difficult to discern the effect of the algorithm choice in the results. Herein, we evaluate the performance of seven modeling algorithms for transferring niche models of Mexican endemic rodents from the mid-20th century to the present and vice versa. We used Mexican endemic rodents as study model because they have restricted and relatively well- known distributions, and because the evolutionary rate of mammals is not high enough to expect climatic niche evolution over this period of time (Martínez-Meyer, Peterson & Hargrove, 2004).

Materials & Methods

Occurrence data of species. We ~~gathered~~ collected occurrence records for the 117 Mexican endemic rodent species (Ramírez-Pulido et al., 2014) from natural history collection databases, journal articles, books, and theses (Supplementary Material, S1 and S2). For each record, we noted the collecting year and grouped occurrences into two time periods to match the climatologies available for the 20th century (see below): 1950-1979 and 1980-2009. We discarded duplicate records to keep only one record per pixel and also discarded those records with questionable taxonomic or geographic certainty. Finally, we retained only those species for which we had at least 10 unique localities per period to produce reliable models (Hernandez et al., 2006). Because some of the algorithms that we analyzed use pseudo-absence data, we randomly selected the same number of unique localities with no presence records as with presence records for each species within the modeling area to generate the pseudo-absence dataset using the *Ecospat* package (Di Cola et al., 2017) in the software R 3.5 (R Core Team, 2017).

Climatic variables. We used 19 bioclimatic variables generated for Mexico by Cuervo-Robayo et al. (2020) for mid-20th century (T1: 1950-1979) and late 20th–early 21st centuries (T2: 1980-2009) (Cuervo-Robayo et al., 2020). These variables were derived from monthly averages of precipitation, minimum and maximum temperatures drawn from climatic stations across Mexico, southern USA, ~~and~~ northern Guatemala and Belize, producing surfaces at 30 arc seconds (~1 km) of spatial resolution, following the same methodology of the Worldclim dataset (Hijmans et al., 2005). The 19 bioclimatic variables summarize the extreme, mean and seasonal patterns of temperature, the ~~and~~ rainfall across the country and have been widely used for ecological analyses and species' distribution modeling (Worldclim has been cited more than 18,000 times, according to Google Scholar). To reduce the number of variables for analyses, we performed a Pearson's correlation analysis for each species to eliminate highly correlated variables (> 0.80) thus reducing model complexity and the risk of overfitting (Radosavljevic & Anderson, 2014; Moreno-Amat et al., 2015; García-Callejas & Araújo, 2016).

Comentado [IR8]: Could you please indicate the species, because in Table S2, there is not 117 species. Is explained in the results but maybe it would be better here.

Comentado [IR9]: What does this mean T?

Comentado [IR10]: idem

Comentado [IR11]: Could you please explain the implication?

Finally, we delimited the area of analysis for each species (i.e., “M” in the BAM framework; (Barve et al., ~~2011)~~2011) by clipping the raster layers to the ecoregions (Olson et al., 2001) in which each species has been recorded, assuming that the boundaries of the ecoregions have represented a barrier to dispersal for the species (Radosavljevic & Anderson, 2014).

Ecological niche modeling. We evaluated the performance of seven popular algorithms for transferring niche models over time: Bioclim (Nix, 1986; Booth et al., 2014), which is a climatic envelope method; two regression-based techniques, namely 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 (Phillips, Anderson & Schapire, 2006), Random Forest (RF; (Breiman, 2001), and Support Vector Machine (SVM; (Drake, Randin & Guisan, 2006). Detailed information about the functioning of each algorithm can be found elsewhere (Guisan, Edwards & Hastie, 2002; Beaumont, Hughes & Poulsen, 2005; Drake, Randin & Guisan, 2006; Elith et al., 2006; Phillips, Anderson & Schapire, 2006; Merow, Smith & Silander, 2013; Booth et al., 2014; Muscarella et al., 2014), but in general, Bioclim describes the niche of a species in terms of an *n*-dimensional quadrangular envelope defined by the range of values in which the species occurs at each individual variable (Nix, 1986). GLM is an extension of linear models without forcing data into unnatural scales and relating the mean of the response variable with the linear combination of the explanatory variables. GAM is a semi-parametric extension of GLM which applies a “smoothing” function to the explanatory variables and replacing the linear predictor with an additive predictor (Guisan, Edwards & Hastie, 2002). BRT is an ensemble method that combines regression trees with a boosting algorithm (Elith, Leathwick & Hastie, 2008). Maxent is an algorithm that uses the maximum entropy principle and a Bayesian procedure to produce a probability surface where entropy is maximized to reflect the environmental suitability of the geographic area for the species (Phillips, Anderson & Schapire, 2006). RF is a combination of individual decision trees that converge in their classification outcome (Prasad, Iverson & Liaw,

Comentado [IR12]: The term elsewhere should be modified, please rewrite the sentence.

Comentado [IR13]: idem

Comentado [IR14]: reference

2006). SVM uses a functional relationship known as kernel to map data onto a new hyperspace (Drake, Randin & Guisan, 2006).

We executed Bioclim in the R package DISMO (Hijmans et al., 2017), where the resulting map is expressed as a continuous model based on the rescaled calculation of percentiles assuming that areas near the median are more suitable. We used the SDM package (Naimi & Araújo, 2016) in R for GLM, GAM, BRT, FR, and SVM. For GLM, we used a binomial response with logit link, quadratic function and AIC for stepwise selection. We implemented GAM with a binomial response with logit link function. We parameterized BRT with a learning rate of 0.005, tree complex of 5 and bag fraction of 0.5. We calibrated RF with 500 trees. We ran Maxent in ENMeval (Muscarella et al., 2014) under default settings but with the clamping and extrapolate options disabled and selected the best model per species with the minimum values of AICc. All output maps were expressed in a continuous scale from 0 to 1.

Comentado [IR15]: or RF (Random Forest)

Comentado [IR16]: All these terms are explained in Table 1, but also should be mentioned at the text.

Comentado [IR17]: Meaning…

Comentado [IR18]: Can you indicate why these parameters, or include references to the works that they previously used them

Comentado [IR19]: AIC is the same as AICc?

For all algorithms, we used 70% of the occurrence records for model calibration and the remaining 30% for validation, and we used the same proportion for pseudo-absences in the algorithms that required so. Then, each resulting model was transformed into a binary map (presence-absence) by selecting a ten-percentile threshold value; i.e., 10% of all occurrence records were left out. We decided to use this threshold to avoid over-prediction due to potentially erroneous occurrences (Radosavljevic & Anderson, 2014). All algorithms were evaluated for their capacity to produce reliable models within time periods via binomial tests to compare model 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 transfer it onto the climatic scenario of period 2, then validate the transference with occurrences from period 2, or calibrate a niche model in period 2 and compare the two maps (Hijmans & Graham, 2006). A problem with these approaches emerges when the number or spatial distribution of occurrences in each period is different, because the differences observed between the resulting maps of the two time periods may have been due to the disparities of the occurrences rather than algorithmic performance. To solve this problem, we followed a cross-temporal approach to identify the species with similar occurrence data structure in the two time periods. First, we calibrated a model in period 1 with occurrences from that period that we named *auto1*, then we generated a second model with the climatic layers of period 1 and the occurrences of period 2, named *cross1*. Second, we calibrated a model using the climatic surfaces of period 2 with the occurrences from that period (*auto2*) and another model with the occurrences of period 1 on the climatic surfaces of period 2 (*cross2*). Third, we converted all resulting maps into binary and geographically compared *auto1* vs. *cross1* and *auto2* vs. *cross2* (see next section). If the overall similarity between the two pairs of maps was <70%, it meant that the difference in the number/distribution of occurrences between time periods was large enough to hinder the response of algorithms to transfer models across temporal scenarios. We repeated this procedure for all species and the seven algorithms and discarded from further analyses those species for which their similarity value was <70%. Species with a dataset suitable for analyses were called *control* and the rest were called *without control*. For *control* species, we transferred models from period 1 to period 2 (forecast) and vice versa (hindcast), and evaluated model performance (fig. 1).

Evaluation of algorithm performance. For each algorithm, every model that was transferred from period 1 to period 2 was compared against the model that was calibrated in period 2, and vice versa, in a pixel-by-pixel fashion, as follows. First, we built a confusion matrix in which the calibration model was the reference; therefore, in the confusion matrix, *a* was the number of presence pixels correctly predicted by the transferred model (sensitivity), *b* was the number of absence pixels erroneously classified as presence (commission error or false positives), *c* was the number of presence pixels erroneously classified as absence (omission error or false negatives), and *d* is the number of absence pixels correctly predicted (specificity). Then, we calculated the following indices: 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), as follows:

FPR is a measure of overestimation, whereas FNR indicates overfitting (Rebelo, Tarroso & Jones, 2010), and both have a scale from 0-1. OI reflects the proportion of overlap between the two maps. Finally, TSS compares the number of correctly predicted pixels minus randomly assigned data within a hypothetical set of perfect predictions. TSS score ranges from -1 to 1, where values between -1 and 0 reflect a model that does not perform better than chance, while a value of 1 is considered a perfect discrimination. Models with TSS values above 0.7 are considered reliable (Allouche, Tsoar & Kadmon, 2006). All calculations were made with the statistical program R 3.5 (R Core Team, 2017).

Statistical Analysis. We performed a Mann-Whitney-Wilcoxon test to evaluate differences between forecast and hindcast transferences (Pohlert, 2016). Also, we implemented a Kruskal-Wallis test to evaluate differences in the transference performance between algorithms, when a difference was found, we carried out a Nemenyi test, which makes pairwise multiple comparisons of mean sums of independent samples. ~~These test~~These tests are implemented in the PMCMR package (Pohlert, 2016). All statistical analyses were performed in R.3.5 (R Core Team, 2017).

Results

Species' occurrences. We found occurrence data for the 117 rodent species endemic to Mexico (Ramírez-Pulido et al., 2014); however, we found sufficient unique records (at least 10) for generating models for each time period only for 44 species (Supplementary Material, S2). There were no sufficient data for 14 species for one of the two periods, whereas there were 59 species for which we did not find 10 unique records for both periods. For the 44 species with sufficient data for further analyses, *Peromyscus melanophrys* (n = 504), *Peromyscus difficilis* (n = 440), *Chaetodipus arenarius* (n = 248), and *Sigmodon mascotensis* (n=191) held the highest number of records.

Comentado [IR20]: I am asking that if you have sufficient records for 44 species maybe it would be better just to work with 44, because if not is confused for the reader.

Comentado [IR21]: idem

Temporal consistency of occurrence data. We evaluated the consistency of occurrence data at each period to perform temporal transferences for all 44 species with sufficient data. We found that the number of species with high model similarity (>70%) between the two time periods was different depending on the algorithm (Table 1). Maxent got the highest number of species, whereas GAM the lowest (Table 1).

Comentado [IR22]: According to table 1, maxent has 44 but also RF, SVM, BRT, GAM…

Comentado [IR23]: In table 1 has also 44

Comentado [IR24]: If this paragraph is related with table 1, should be reanalyzed.

Niche models and model transferences. In general, all algorithms produced reliable models for calibration in both time periods. Binomial test for models calibrated in the two climatic scenarios were significantly different from random expectations for all species, except for GLM models for *Callospermophilis madrensis*, *Dasyprocta mexicana* and *Dipodomys phillipsii* for T1 (1959-1979), and *Dipodomys phillipsii, Neotamias durangae*, *Noetoma goldmani* and *Oryzomys guerrerensis* for T2 (1979-2009); a BRT model for *Pappogeomys bulleri* for T2; and a Maxent model for *Dasyprocta mexicana* for T2 (Supplementary Material, S3?).

Comentado [IR25]: You have more than one table in the SP, o please indicate to which one.

We observed differences in algorithm performance when models were transferred across climatic scenarios (fig. 2). Some algorithms performed differently for forecasting and hindcasting: our results showed that Bioclim, GAM and Maxent generally produced higher TSS values for hindcasting than for forecasting, with significant differences between directions only for Maxent (w = 139, *p* = 0.04). Similar results were observed for the omission index (OI), with significant differences between directions for Maxent (w = 140, *p* = 0.043).

Comentado [IR26]: Is not explained in the text

We found significant differences between algorithms for transferring models from the present to the past scenario ($H = 17.71$, $p = 0.006$) and from the past to the present ($H = 21.50$, $p < 0.01$), being BRT the algorithm with the highest mean TSS value ($\bar{x} = 0.84 \pm 0.08$) and Bioclim the lowest ($\bar{x} = 0.46 \pm 0.24$) in forecasting. For hindcasting, Maxent obtained the highest mean value ($\bar{x} = 0.82 \pm 0.11$) and Bioclim the lowest ($\bar{x} = 0.60 \pm 0.15$). We observed a similar pattern for OI with significant differences between algorithms for hindcasting ($H = 27.27$, $p = 0.014$) and forecasting ($H = 15.93$, $p < 0.01$). We observed the greatest variation in forecasting for Bioclim ($\bar{x} = 0.46 \pm 0.24$) and GAM ($\bar{x} = 0.54 \pm 0.40$) and, on the contrary, RF showed a very consistent behavior and little variation in both directions (forecasting: $\bar{x} = 0.81 \pm 0.09$; hindcasting: $\bar{x} = 0.80 \pm 0.075$), and GAM for hindcasting ($\bar{x} = 0.8 \pm 0.07$).

The false negative rate (FNR), which indicates model overestimation, showed significant differences between algorithms for hindcasting ($H = 15.93$, $p = 0.014$) and forecasting ($H = 27.27$, $p < 0.01$). Bioclim showed a difference with BRT ($p = <0.01$) and RF ($p = 0.03$) for forecasting and Maxent ($p = <0.01$) for hindcasting. GAM showed the highest FNR for forecasting ($\bar{x} = 0.29 \pm 0.39$) and GLM for hindcasting ($\bar{x} = 0.10 \pm 0.06$), both algorithms with the highest variation in each direction. In turn, the false positive rate (FPR; an indication of overfitting) was not different for hindcasting ($H = 10.96$, $p = 0.089$), but it was so for forecasting ($H = 13.22$, $p=0.039$). We observed a difference in FPR between directions only for Maxent ($w = 301$, $df = 6$, $p = 0.04$). Conversely, GAM showed the highest FPR values for forecasting.

Discussion

Transferences of ecological niche models over time have been widely used to assess the response of biodiversity to climatic changes (Allouche, Tsoar & Kadmon, 2006; Araújo, Thuiller & Pearson, 2006; Rebelo, Tarroso & Jones, 2010; Peterson et al., 2015a). Similar to our results, the variation between algorithms in niche model transfers to alternative climatic scenarios seems to be the rule and has been observed in multiple comparative studies

(Moreno-Amat et al., 2015; Beaumont et al., 2016; Bell & Schlaepfer, 2016; García-Callejas & Araújo, 2016; Liang et al., 2018; Qiao et al., 2019). However, to our knowledge, no previous studies have analyzed the effect of occurrence data structure and direction of transfers in the performance of algorithms.

In this study, we evaluated the performance and consistency of seven popular algorithms for transferring niche models to past (hindcast) and future (from the past to the present; forecast) climatic conditions for Mexican endemic rodents as study model. We decided to use endemic species for two reasons: (1) the ecological niche of the species is better reconstructed for species with restricted ecological conditions than for those that occupy a great diversity of environments (Raxworthy et al., 2007); and (2) the climatic databases used in this analysis were developed specifically for Mexico (Cuervo-Robayo et al., 2020). Our analyses showed three main results: first, the quality of input data and the algorithm had a significant effect on model output and performance. Second, algorithms did not perform equally for hindcasting and forecasting, where we found generally higher performance for the latter. And third, we found that algorithms performed differently for model transferences, being Random Forest (RF), Boosted Regression Trees (BTR) and Maxent consistently robust, whereas Bioclim showed the lowest performance.

Comentado [IR27]: You should discuss your results

Algorithm performance for model transferences has been analyzed under different approaches, including the use of virtual species (García-Callejas & Araújo, 2016) and real species of different biological groups (Heikkinen, Marmion & Luoto, 2012; Beaumont et al., 2016; Fan et al., 2018), as well as testing different sets of variables (Barbet-Massin & Jetz, 2014; Warren et al., 2014) and climate change projections to the past and future (Dobrowski et al., 2011; Veloz et al., 2012). In this study, we assessed algorithm performance for a single taxonomic group under a retrospective approach and found that occurrence data structure between time periods plays a critical role. We know that an adequate sampling is necessary to make a good characterization of the species’ ecological niche in the calibration scenario before transferring a model to different scenarios (Araújo, Pearson & Rahbek, 2005), but this is the first time that disparities in the occurrence data between time periods are evaluated in the context of climate change analyses.

Comentado [IR28]: Please indicate the benefits of this method

One important issue when using correlative modeling under a retrospective approach to evaluate the geographic responses of species to climate change is to decouple the effect of data bias or incompleteness from the signal of geographic change. A logical way to proceed when you have two temporally distinct sets of occurrence data and climatic scenarios (i.e., T1 and T2) is to calibrate a model in T1 with the occurrences and climatic scenario from that time period, transfer the resulting model to T2, and use the occurrences of T2 to validate the transfer, and vice versa. However, this approach may have a problem of inconsistency of the occurrences between the two time periods, ergo, if the quantity or quality of occurrence data from the two periods are different, then the two niche models will result somewhat different and transferences will be affected in an unknown manner. In such case, the geographic shift observed under these conditions is mostly due to the occurrence differences, rather than to the responses of species to climatic changes. To avoid this pitfall, our initial step was to test for data consistency between the two time periods by carrying out a cross-modeling procedure. To do so for the 44 species with sufficient data in the two time periods, we first calibrated a model with the climatic scenario and occurrences form T1, and then, we calibrated a second model with occurrences from T2 under the climatic scenario of T1 and compared the two resulting maps, and vice versa. If the data from the two time periods were consistent, the two maps should result highly similar, otherwise they would be significantly different. Interestingly, we found that the number of species with consistent data was different depending on the algorithm, being Maxent the most robust (n = 21), whereas the General Additive Model (GAM) obtained the lowest number (n = 3) (Table 1). This result corroborates the fact that algorithms have different sensitivity to data bias or incompleteness (Moreno-Amat et al., 2015); therefore, if the user knows or suspects about this problem in the dataset, a safer decision is to use a low data-demanding algorithm (e.g., Maxent, Bioclim).

There are at least three assumptions when transferring niche models over time: (1) the species is in equilibrium with the environment (Araújo, Pearson & Rahbek, 2005), (2) the species maintains the same response to environmental conditions between time periods (Anderson 2011), and (3) the ecological niche remains conserved along time periods (Soberón & Nakamura, 2009; Wiens et al., 2009; Sequeira et al., 2018). However, other factors have a strong effect on the capacity of niche models to transfer over time, the most critical one being the variability produced by at least three sources: (1) data completeness and structure (i.e., quality and quantity) (García-Callejas & Araújo, 2016; García- et al., 2023); (2) algorithm performance (Heikkinen, Marmion & Luoto, 2012; Beaumont et al., 2016; Liang et al., 2018; Qiao et al., 2019); and (3) taxonomic idiosyncrasies (Sequeira et al., 2018; Regos et al., 2019). Of the seven algorithms that we tested, Random Forest (RF) showed the greatest predictive capacity and lowest overprediction. This algorithm has proved robust for transferences using virtual species (García-Callejas & Araújo, 2016), good performance for interpolations (Liang et al., 2018), relatively low overprediction (Mi et al., 2017), and high performance (Bell & Schlaepfer, 2016). Our results are consistent with other studies that found BRT having a good performance (García-Callejas & Araújo, 2016; Liang et al., 2018; Sequeira et al., 2018) and high variability (Bell & Schlaepfer, 2016). Maxent was also an algorithm that showed a high performance, particularly for hindcasting. Maxent is the most widely used algorithm (Radosavljevic & Anderson, 2014; Moreno-Amat et al., 2015) in part because it is robust to data biases (Heikkinen, Marmion & Luoto, 2012; Liang et al., 2018; Sequeira et al., 2018). In turn, Bioclim presented the lowest performance, with low overfitting but high overprediction and variability between projections. Bioclim is sensitive to outliers or insufficient sampling, which might be the reason for its poor performance (Dobrowski et al., 2011; Beaumont et al., 2016).

Another important and novel result derived from our analyses is that most algorithms performed differently for transferring models to future scenarios than to the past. Bioclim and GAM performed better for hindcasting than for forecasting, whereas RF, Support Vector Machine (SVM) and BRT behaved the opposite. Maxent and the Generalized Linear Model (GLM) were consistent for the two time periods. These differences may reflect the capacity of models to deal with non-analog climates (Pearson et al., 2006). For instance, Bioclim is sensitive to extreme values, so it may be more affected by larger differences between climatic scenarios, as compared to Maxent that has the capacity to clamp and extrapolate into non-analog values (Phillips & Dudík, 2008). More importantly, these results ~~suggests~~suggest that the choice of an algorithm has implications beyond its simple capacity to produce adequate models in the calibration scenario.

Comentado [IR29]: As you are in the discussion your results should be compared with the previous ones. If not you can merge results and discussion sections to minimize redundancy.

Finally, a potential limitation of this study is that the taxonomy of Mexican rodents is continuously updated. While our research relied on the latest revision of Mexican mammals (Ramírez-Pulido et al. 2014~~),recent~~), 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). As a result, the quantity and spatial structure of occurrences for these species may undergo some modifications. Nonetheless, based on our calibration tests, we believe that the generality of our findings remains unaffected by these taxonomic changes

Finally, when the aim of a study is to explore the potential distributional responses of species to climate change under a niche modeling approach, we recommend the following: First, perform a thorough screening and cleaning of occurrence data and identify the most important variables to characterize the ecological niche of the focal species. Then, perform a cross-modeling procedure between time periods in a retrospective fashion (past to present and vice versa) using different algorithms to identify the effect of sampling bias or incompleteness on model transferences, and decide whether the target species is suitable for the analysis. Split the occurrence data into two time periods (e.g. present-near, past) and model between times with different algorithms or perform a consensus to identify the model with the highest transferability over time according to the specific data structure. Finally, identify FNR and FPR to ~~make a decision~~decide on the error that best fits the question to be answered by the transfer.

Conclusions

Comentado [IR30]: Please, indicate more conclusions of the work.

We found that both the quantity and quality of data, along with the choice of the modeling algorithm significantly influence model transferability. Furthermore, we observed that algorithms exhibit unequal performance when transferring models from the past to the present compared to the reverse process. Consequently, we assert that not all species are suitable candidates for transferring models to alternative climatic scenarios; only those species with sufficient occurrence records to produce robust niche models, should be considered for projections time periods.

Comentado [IR31]: Which data is needed?

Additionally, not all algorithms perform uniformly across different occurrence data structure and levels of completeness. Some algorithms are more sensitive to small sample sizes and non-analog climates. Therefore, we recommend conducting preliminary tests on algorithmic performance to specific occurrence datasets in a retrospective manner before proceeding with model transferences to future climatic scenarios.

Acknowledgements

This contribution is a fulfillment of the Graduate Doctoral Degree program in Biological Sciences of the Universidad Nacional Autónoma de México (UNAM) to CNM-A, who acknowledges the Posgrado en Ciencias Biológicas, Instituto de Biología-UNAM for the logistic support and the financial support provided by the Consejo Nacional de Humanidades, Ciencias y Tecnologías of Mexico (CONAHCYT) through the scholarship 211508. We thank Livia León-Paniagua and

Fausto Mendez for her guidance during the development of the study and Town Peterson for his valuable comments to the manuscript.

References

Comentado [IR32]: Check the 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 AL, González-Cózatl FX, Engstrom MD, Rogers DS. 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 RP, Lew D, Peterson AT. 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, Pearson R, Rahbek C. 2005. Equilibrium of species' distribution with climate. *Ecography* 28:693–695.

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

Araújo M, Thuiller W, Pearson RG. 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.

Araújo MB, Whittaker RJ, Ladle RJ, Erhard M. 2005. Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography* 14:529–538. DOI: 10.1111/j.1466-822X.2005.00182.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 CH, 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 SP, Peterson AT, 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 LJ, Gallagher R V, Thuiller W, Downey PO, Leishman MR, Hughes L. 2009. Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. :409–420. DOI: 10.1111/j.1472-4642.2008.00547.x.

Beaumont LJ, Graham E, Duursma DE, Wilson PD, Cabrelli A, Baumgartner JB, Hallgren W, Esperón-Rodríguez M, Nipperess DA, Warren DL, Laffan SW, 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: 10.1016/j.ecolmodel.2016.10.004.

Beaumont LJ, Hughes L. 2002. Potential changes in the distributions of latitudinally restricted Australian butterfly species in response to climate change. *Global Change Biology* 8:954–971. DOI: 10.1046/j.1365-2486.2002.00490.x.

Beaumont LJ, 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 DM, Schlaepfer DR. 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.

Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F. 2012. Impacts of climate change on the future of biodiversity. *Ecology letters*:365–377. DOI: 10.1111/j.1461-0248.2011.01736.x.

Booth TH, Nix H a., Busby JR, Hutchinson MF. 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 RD, Ordóñez-Garza N, Thompson CW, Wright EA, Ceballos G, Kilpatrick CW, Schmidly DJ. 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 MB, Hirzel AH. 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27:437–448.

Di Cola V, Broennimann O, Petitpierre B, Breiner FT, D'Amen M, Randin C, Engler R, Pottier J, Pio D, Dubuis A, Pellissier L, Mateo RG, 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.

Cruz-Gómez A, Castro-Campillo A, Ávila-Valle ZA, 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 MA, Meneses Mosquera AK, 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.

Dobrowski SZ, Thorne JH, Greenberg JA, Safford HD, Mynsberge AR, Crimmins SM, Swanson AK. 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 JM, 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 CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JMC, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151.

Elith J, Leathwick JR, 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 DM, Rumpf SB, 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 JY, Zhao NX, Li M, Gao WF, Wang ML, Zhu GP. 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 AH, Bell JF. 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 MB. 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 RA, Cabeza M, Altwegg R, Araújo MB. 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 TC, 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 MS, Din BH, Tan SH, 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 RK, 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 PA, Graham CH, Master LL, Albert DL. 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 RJ, Cameron SE, Parra JL, Jones PG, 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 RJ, 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 ARJ, Phillips S, Leathwick J, Elith J, Hijmans MRJ. 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, Lobo JM, Hortal J. 2008. Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions* 14:885–890. DOI: 10.1111/j.1472-4642.2008.00496.x.

Kharouba HM, Algar AC, Kerr JT. 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 ST, Pérez-Consuegra SG, Patton JL. 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 SA, Rios Dominguez AF, de Valpine P, Beissinger SR. 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 NE, 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 AT, Hargrove WW. 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 MJ, Edwards Jr TC, Guisan A, McMahon SM, Normand S, Thuiller W, Wüest RO, Zimmermann NE, 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 MJ, 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 RG, Nieto-Lugilde D, Morueta-Holme N, Svenning JC, 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.

Moritz C, Patton JL, Conroy CJ, Parra JL, White GC, Beissinger SR. 2008. Impact of a century of climate change on small mammal communities in Yosemite national park, USA. *Science* 322:261–264. DOI: 10.1126/science.1163428.

Muscarella R, Galante PJ, Soley-Guardia M, Boria RA, Kass JM, Uriarte M, Anderson RP. 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. DOI: 10.1111/2041-210X.12261.

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

Nix HA. 1986. *A biogeographic analysis of Australian Elapid snakes. In: Longmore, R.* Canberra.

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'amico J a., Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR. 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.

Pacifici M, Foden WB, Visconti P, Watson JEM, Butchart SHM, Kovacs KM, Scheffers BR, Hole DG, Martin TG, Akçakaya HR, Corlett RT, Huntley B, Bickford D, Carr JA, Hoffmann AA, Midgley GF, Pearce-Kelly P, Pearson RG, Williams SE, Willis SG, 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 RG. 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 RG, 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 AT. 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 AT, 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 AT, Navarro-Siguenza AG, Martinez-Meyer E, Cuervo-Robayo AP, Berlanga H, Soberón J. 2015a. Twentieth century turnover of Mexican endemic avifaunas: Landscape change versus climate drivers. *Science Advances* 1:e1400071–e1400071. DOI: 10.1126/sciadv.1400071.

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

Phillips SJ. 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 SJ, Anderson RP, Schapire RE. 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 JA, 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 AM, Iverson LR, 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, Feng X, Escobar LE, Peterson AT, 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 RP. 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 AL, Arroyo-Cabrales J. 2014. *List of recent land mammals of Mexcio*. Lubbock, Texas. DOI: 10.1177/002194368302000404.

Raxworthy CJ, Ingram CM, Rabibisoa N, Pearson RG. 2007. Applications of ecological niche modeling for species delimitation: A review and empirical evaluation using day geckos (Phelsuma) from Madagascar. *Systematic Biology* 56:907–923. DOI: 10.1080/10635150701775111.

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 JP, 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 EM, Monahan WB, Parra JL, Cameron SE, Brashares JS. 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 AMM, Bouchet PJ, Yates KL, Mengersen K, Caley MJ. 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. 2011. What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling* 222:1343–1346. DOI: 10.1016/j.ecolmodel.2011.01.018.

Sillero N, Campos JC, Arenas-Castro S, Barbosa AM. 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.

Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BFN, De Siqueira MF, Grainger A, Hannah L, Hughes L, Huntley B, Van Jaarsveld AS, Midgley GF, Miles L, Ortega-Huerta M a, Peterson AT, Phillips OL, Williams SE. 2004. Extinction risk from climate change. *Nature* 427:145–8. DOI: 10.1038/nature02121.

Vaz UL, Cunha HF, Nabout JC. 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.

Veloz SD, Williams JW, Blois JL, He F, Otto-Bliesner B, Liu Z. 2012. No-analog climates and shifting realized niches during the late quaternary: Implications for 21st-century predictions by species distribution models. *Global Change Biology* 18:1698–1713. DOI: 10.1111/j.1365-2486.2011.02635.x.

Waltari E, Guralnick RP. 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.

Warren DL, Wright AN, Seifert SN, Shaffer HB. 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 WT. 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 JA, Stralberg D, Jongsomjit D, Howell CA, Snyder MA. 2009. Niches, models, and climate change: assessing the assumptions and uncertainties. *PNAS* 106 Suppl:19729–19736. DOI: 10.1073/pnas.0901639106.

Yates KL, Bouchet PJ, Caley MJ, Mengersen K, Randin CF, Parnell S, Fielding AH, Bamford AJ, Ban S, Barbosa AM, Dormann CF, Elith J, Embling CB, Ervin GN, Fisher R, Gould S, Graf RF, Gregr EJ, Halpin PN, Heikkinen RK, Heinänen S, Jones AR, Krishnakumar PK, Lauria V, Lozano-Montes H, Mannocci L, Mellin C, Mesgaran MB, Moreno-Amat E, Mormede S, Novaczek E, Oppel S, Ortuño Crespo G, Peterson AT, Rapacciuolo G, Roberts JJ, Ross RE, Scales KL, Schoeman D, Snelgrove P, Sundblad G, Thuiller W, Torres LG, Verbruggen H, Wang L, Wenger S, Whittingham MJ, Zharikov Y, Zurell D, Sequeira AMM. 2018. Outstanding Challenges in the Transferability of Ecological Models.