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

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

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 Department of Zoology, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, Mexico

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

Ecological niche modeling (ENM) is an approach to infer 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 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 do so, we followed a retrospective approach by transferring models from the near past (1950-1979) to the present (1980-2009) and vice versa. We 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 test different algorithms *prior* transferring models to future scenarios under a retrospective approach.

1 A retrospective approach for evaluating ecological

2 niche modeling transferences over time: The case of

3 Mexican endemic rodents

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

5

6 [1] Posgrado de Ciencias Biológicas, Universidad Nacional Autónoma de México, Mexico City
7 04510, Mexico.
8 [2] Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico.
9 [3] Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico.

10

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

15

16 Abstract

17 Ecological niche modeling (ENM) is an approach to infer the suitable conditions for species

18 persistence and their potential geographic distributions. ENM is extensively used to assess the

19 potential effects of climate change on species' distributions, although modeling algorithms are

20 acknowledged as an important source of uncertainty in climate change projections. A problem is

21 that these models cannot be properly assessed in the context of future projections. In this study,

22 we evaluated the performance of seven popular modeling algorithms (Bioclim, Generalized

23 Additive Models (GAM), Generalized Linear Models (GLM), Boosted Regression Trees (BRT),

24 Maxent, Random Forest (RF), and Support Vector Machine (SVM)) for transferring ENM over

25 time for Mexican endemic rodents. To do so, we followed a retrospective approach by

26 transferring models from the near past (1950-1979) to the present (1980-2009) and vice versa.

27 We found three important results: (1) The quality of input data and the algorithm had a significant

28 effect on model output and performance; (2) algorithm performance was different for transferring

29 models to the future than to the past; and (3) the most robust algorithms were RF, BRT and

30 Maxent, whereas Bioclim was the least consistent. In conclusion, algorithm choice is critical for

Comment (margin): This first sentence could be:

Ecological niche modeling allows us to infer suitable conditions for persistence and estimate geographic distributions of species.

31 ~~transferring ENM over time.~~ Since no algorithm performed consistently better than the rest, we

32 recommend ~~to test~~ different algorithms *prior* ~~transferring~~ models to future scenarios under a

33 retrospective approach.

34

> But, you just said that Bioclim was worst!

Introduction

36 ~~Current~~ climate change is affecting biodiversity in ~~a myriad of~~ ways (Parmesan &

37 Yohe, 2003; Bellard et al., 2012; MacLean et al., 2018; Widick & Bean, 2019; Cuervo-Robayo et

38 al., 2020; Habibullah et al., 2022) and biotas are changing due to ~~alterations~~ in ~~the~~ abundance

39 and ~~distributional shifts~~ of species (Moritz et al., 2008; Kharouba, Algar & Kerr, 2009; Dobrowski

40 et al., 2011; Pacifici et al., 2015; Peterson et al., 2015b; Jaroszynska et al., 2023). ~~For some~~

41 ~~years now,~~ ecological niche modeling (ENM) is a popular tool for analyzing geographic

42 distributions of species in the context of past and ~~current climatic~~ changes (Araújo et al., 2005;

43 Beaumont et al., 2009; García-Callejas & Araújo, 2016; Sequeira et al., 2018; Yates et al., 2018;

44 García- et al., 2023). ENM includes a suite of concepts and methodologies based on ecological

45 niche theory ~~that has taken advantage of~~ other disciplines, such as informatics, geographic

46 information systems (GIS), and statistics (Austin, 2007; Jiménez-Valverde, Lobo & Hortal, 2008;

47 Soberón & Nakamura, 2009; Wiens et al., 2009).

48 ENMs are correlative methods that require georeferenced species~~'~~ occurrence data

49 and ~~a set of~~ predictor variables in the form of GIS raster layers to reconstruct the ecological

50 characteristics that define the conditions ~~where species inhabit (i.e., its ecological niche) and~~

51 then project these conditions onto the landscape to produce a geographical expression of ~~the~~

52 ~~species'~~ niche, commonly referred to as its potential distribution (Elith et al., 2006; Peterson,

53 2011; Sillero, 2011). Several methods ~~have been implemented~~ to produce niche models, which

54 can be classified ~~according to~~ the type of input occurrence data as presence-only (e.g., Bioclim,

55 environmental distances), presence-absence (e.g., Generalized Linear Models [GLM],

56 Generalized Additive Models [GAM]), presence-pseudoabsence (e.g., Genetic Algorithm for

57 Rule-set Prediction [GARP]), and presence-background (e.g., Ecological Niche Factor Analysis

58 [ENFA], Maxent) (Phillips, 2008; Barbet-Massin et al., 2012; Warton & Aarts, 2013; Fan et al.,

59 2018; Qiao et al., 2019; Sillero et al., 2023). From these, presence-absence methods are

60 generally more robust when unequivocal absence data exist (Brotons et al., 2004; Golicher et

61 al., 2012; Merow et al., 2014); however, in most cases, reliable absences are lacking. Therefore,

62 presence-only, presence-pseudoabsence and, presence-background algorithms have become

63 popular in the last two decades (Vaz, Cunha & Nabout, 2015).

64   ENM has been broadly used to evaluate the potential geographic consequences of

65 climatic changes by projecting the modeled species' niche under current climatic conditions to

66 past or future climatic scenarios, a process also known as model transference (Pearson &

67 Dawson, 2003; Peterson, Martínez-Meyer & González-Salazar, 2004; Thomas et al., 2004;

68 Waltari & Guralnick, 2009; Heikkinen, Marmion & Luoto, 2012; Garcia et al., 2016). At least three

69 main assumptions are behind model transferences over time that may or may not hold true: (1)

70 the ecological niche does not evolve during the timeframe when the transference is performed

71 (Soberón & Nakamura, 2009); (2) during a climatic change, even if the interactions among

72 variables change, the relationship of those variables with the species remains constant (Hijmans

73 & Graham, 2006; Wiens et al., 2009); and (3) the species is in equilibrium with the environment

74 in the calibration time/area, meaning that they occupy the available suitable areas that are

75 accessible to them (Sequeira et al., 2018; Yates et al., 2018).

76   Furthermore, a key issue in model transference is the effect of non-analog climates

77 (i.e., the set of environmental conditions that exist in an alternative climatic scenario different

78 from the calibration one) on model performance (Sequeira et al., 2018). Non-analog climates

79 impose a challenge to algorithms because no calibration exists for those climates, thus, when

80 algorithms face non-analogues in alternative climatic scenarios (past or future), algorithms

81 respond differently according to the instructions of the program, producing contrasting or even

82 contradicting results(Pearson, 2006; Araújo & Peterson, 2012; Essl et al., 2023). This is

83 particularly problematic in future projections because there are no occurrence data that can be

84 used to evaluate algorithm performance. Thus, one solution to evaluate algorithm performance

85 in climate change studies is to project models between two time periods ~~for which they have~~

86 occurrence data, for example, from the relatively recent past to the present (Rubidge et al.,

87 2011; Piirainen et al., 2023).

88 ~~Several studies have focused on evaluating~~ <mark>the reliability of niche models for a</mark>

89 <mark>specific time</mark> (Beaumont & Hughes, 2002; Araújo, Thuiller & Pearson, 2006; Maiorano et al.,

90 2011; Werkowska et al., 2017; Sequeira et al., 2018), but ~~only a few have analyzed the~~

91 performance of algorithms for transferring models over time (Prasad, Iverson & Liaw, 2006;

92 Hijmans & Graham, 2006; Kharouba, Algar & Kerr, 2009; Rubidge et al., 2011; Moreno-Amat et

93 al., 2015; García-Callejas & Araújo, 2016), with performance results spanning from good to bad

94 for different species and algorithms, making it difficult to discern the effect of the algorithm

95 choice in the results. Herein, we evaluate the performance of seven modeling algorithms for

96 transferring niche models of Mexican endemic rodents from the mid-20th century to the present

97 and vice versa. We used Mexican endemic rodents ~~as study model~~ because they have restricted

98 and relatively well- known distributions and because ~~the evolutionary rate of mammals is not~~

99 ~~high enough to~~ expect climatic niche evolution over this period of time (Martínez-Meyer,

100 Peterson & Hargrove, 2004).

101

102 **Materials & Methods**

103 *Occurrence data of species.* We gathered occurrence records for the 117 Mexican endemic

104 rodent species (Ramírez-Pulido et al., 2014) from natural history collection databases, journal

105 articles, books, and theses (Supplementary Material). For each record, we noted the collecting

106 year and grouped occurrences into two time periods to match the climatologies available for the

107 20th century (see below): 1950-1979 and 1980-2009. We discarded duplicate records to keep

108 only one record per pixel and also discarded those records with questionable taxonomic or

109 geographic certainty. Finally, we retained only those species for which we had at least 10 unique

110 localities per period to produce reliable models (Hernandez et al., 2006). Because some of the

111 algorithms that we analyzed use pseudo-absence data, we randomly selected the same number

112   of unique localities with no presence records as with presence records for each species within

113   the modeling area to generate the pseudo-absence dataset using the *Ecospat* package (Di Cola

114   et al., 2017) in the software R 3.5 (R Core Team, 2017).

115   *Climatic variables.* We used 19 bioclimatic variables generated for Mexico by Cuervo-

116   Robayo et al. (2020) for mid-20th century (T1: 1950-1979) and late 20th–early 21st centuries

117   (T2: 1980-2009) (Cuervo-Robayo et al., 2020). These variables were derived from monthly

118   averages of precipitation, minimum and maximum temperatures drawn from climatic stations

119   across Mexico, southern USA and northern Guatemala and Belize, producing surfaces at 30 arc

120   seconds (~1 km) of spatial resolution, following the same methodology of the Worldclim dataset

121   (Hijmans et al., 2005). The 19 bioclimatic variables summarize the extreme, mean and seasonal

122   patterns of temperature and rainfall across the country and have been widely used for ecological

123   analyses and species' distribution modeling (Worldclim has been cited more than 18,000 times,

124   according to Google Scholar). To reduce the number of variables for analyses, we performed a

125   Pearson's correlation analysis for each species to eliminate highly correlated variables (> 0.80)

126   thus reducing model complexity and the risk of overfitting (Radosavljevic & Anderson, 2014;

127   Moreno-Amat et al., 2015; García-Callejas & Araújo, 2016). Finally, we delimited the area of

128   analysis for each species (i.e., "M" in the BAM framework; (Barve et al., 2011) by clipping the

129   raster layers to the ecoregions (Olson et al., 2001) in which each species has been recorded,

130   assuming that the boundaries of the ecoregions have represented a barrier to dispersal for the

131   species (Radosavljevic & Anderson, 2014).

132   *Ecological niche modeling.* We evaluated the performance of seven popular algorithms

133   for transferring niche models over time: Bioclim (Nix, 1986; Booth et al., 2014), which is a

134   climatic envelope method; two regression-based techniques, namely Generalized Additive

135   Models (GAM) and Generalized Linear Models (GLM) (Guisan, Edwards & Hastie, 2002); and

136   four machine learning algorithms: Boosted Regression Trees (BRT; (Elith, Leathwick & Hastie,

137   2008), Maxent (Phillips, Anderson & Schapire, 2006), Random Forest (RF; (Breiman, 2001), and

138   Support Vector Machine (SVM; (Drake, Randin & Guisan, 2006). Detailed information about the

139 functioning of each algorithm can be found elsewhere (Guisan, Edwards & Hastie, 2002;

140 Beaumont, Hughes & Poulsen, 2005; Drake, Randin & Guisan, 2006; Elith et al., 2006; Phillips,

141 Anderson & Schapire, 2006; Merow, Smith & Silander, 2013; Booth et al., 2014; Muscarella et

142 al., 2014), but in general, Bioclim describes the niche of a species in terms of an *n*-dimensional

143 quadrangular envelope defined by the range of values in which the species occurs at each

144 individual variable (Nix, 1986). GLM is an extension of linear models without forcing data into

145 unnatural scales and relating the mean of the response variable with the linear combination of

146 the explanatory variables. GAM is a semi-parametric extension of GLM which applies a

147 "smoothing" function to the explanatory variables and replacing the linear predictor with an

148 additive predictor (Guisan, Edwards & Hastie, 2002). BRT is an ensemble method that combines

149 regression trees with a boosting algorithm (Elith, Leathwick & Hastie, 2008). Maxent is an

150 algorithm that uses the maximum entropy principle and a Bayesian procedure to produce a

151 probability surface where entropy is maximized to reflect the environmental suitability of the

152 geographic area for the species (Phillips, Anderson & Schapire, 2006). RF is a combination of

153 individual decision trees that converge in their classification outcome (Prasad, Iverson & Liaw,

154 2006). SVM uses a functional relationship known as kernel to map data onto a new hyperspace

155 (Drake, Randin & Guisan, 2006).

156   We executed Bioclim in the R package DISMO (Hijmans et al., 2017), where the

157 resulting map is expressed as a continuous model based on the rescaled calculation of

158 percentiles assuming that areas near the median are more suitable. We used the SDM package

159 (Naimi & Araújo, 2016) in R for GLM, GAM, BRT, FR, and SVM. For GLM, we used a binomial

160 response with logit link, quadratic function and AIC for stepwise selection. We implemented

161 GAM with a binomial response with logit link function. We parameterized BRT with a learning

162 rate of 0.005, tree complex of 5 and bag fraction of 0.5. We calibrated RF with 500 trees. We ran

163 Maxent in ENMeval (Muscarella et al., 2014) under default settings but with the clamping and

164 extrapolate options disabled and selected the best model per species with the minimum values

165 of AICc. All output maps were expressed in a continuous scale from 0 to 1.

166   For all algorithms, we used 70% of the occurrence records for model calibration and

167   the remaining 30% for validation, and we used the same proportion for pseudo-absences in the

168   algorithms that required so. Then, each resulting model was transformed into a binary map

169   (presence-absence) by selecting a ten-percentile threshold value; i.e., 10% of all occurrence

170   records were left out. We decided to use this threshold to avoid over-prediction due to potentially

171   erroneous occurrences (Radosavljevic & Anderson, 2014). All algorithms were evaluated for

172   their capacity to produce reliable models within time periods via binomial tests to compare model

173   results against random expectations (Anderson, Lew & Peterson, 2003).

174   *Model transferences.* The logical procedure to evaluate the capacity of algorithms to

175   transfer niche models across different temporal scenarios involves calibrating a niche model in

176   period 1 and transfer it onto the climatic scenario of period 2, then validate the transference with

177   occurrences from period 2, or calibrate a niche model in period 2 and compare the two maps

178   (Hijmans & Graham, 2006). A problem with these approaches emerges when the number or

179   spatial distribution of occurrences in each period is different, because the differences observed

180   between the resulting maps of the two time periods may have been due to the disparities of the

181   occurrences rather than algorithmic performance. To solve this problem, we followed a cross-

182   temporal approach to identify the species with similar occurrence data structure in the two time

183   periods. First, we calibrated a model in period 1 with occurrences from that period that we

184   named *auto1*, then we generated a second model with the climatic layers of period 1 and the

185   occurrences of period 2, named *cross1*. Second, we calibrated a model using the climatic

186   surfaces of period 2 with the occurrences from that period (*auto2*) and another model with the

187   occurrences of period 1 on the climatic surfaces of period 2 (*cross2*). Third, we converted all

188   resulting maps into binary and geographically compared *auto1* vs. *cross1* and *auto2* vs. *cross2*

189   (see next section). If the overall similarity between the two pairs of maps was <70%, it meant

190   that the difference in the number/distribution of occurrences between time periods was large

191   enough to hinder the response of algorithms to transfer models across temporal scenarios. We

192   repeated this procedure for all species and the seven algorithms and discarded from further

193 analyses those species for which their similarity value was <70%. Species with a dataset

194 suitable for analyses were called *control* and the rest were called *without control*. For *control*

195 species, we transferred models from period 1 to period 2 (forecast) and vice versa (hindcast),

196 and evaluated model performance (fig. 1).

197 *Evaluation of algorithm performance.* For each algorithm, every model that was

198 transferred from period 1 to period 2 was compared against the model that was calibrated in

199 period 2, and vice versa, in a pixel-by-pixel fashion, as follows. First, we built a confusion matrix

200 in which the calibration model was the reference; therefore, in the confusion matrix, *a* was the

201 number of presence pixels correctly predicted by the transferred model (sensitivity), *b* was the

202 number of absence pixels erroneously classified as presence (commission error or false

203 positives), *c* was the number of presence pixels erroneously classified as absence (omission

204 error or false negatives), and *d* is the number of absence pixels correctly predicted (specificity).

205 Then, we calculated the following indices: True Skill Statistics (TSS; Equation 1), Overlap Index

206 (OI, Equation 2), False Negative Rate (FNR; Equation 3) and False Positive Rate (FPR;

207 Equation 4) (Fielding & Bell, 1997), as follows:

208

209
$$TSS = \frac{a}{a+c} + \frac{d}{b+d} - 1 \; Eq.1$$

210
$$OI = a/(a+c) \; Eq.2$$

211
$$FNR = \frac{b}{b+d} \; Eq.3$$

212
$$FPR = \frac{c}{a+c} \; Eq.4$$

213

214 FPR is a measure of overestimation, whereas FNR indicates overfitting (Rebelo,

215 Tarroso & Jones, 2010), and both have a scale from 0-1. OI reflects the proportion of overlap

216 between the two maps. Finally, TSS compares the number of correctly predicted pixels minus

217 randomly assigned data within a hypothetical set of perfect predictions. TSS score ranges from -

218   1 to 1, where values between -1 and 0 reflect a model that does not perform better than chance,

219   while a value of 1 is considered a perfect discrimination. Models with TSS values above 0.7 are

220   considered reliable (Allouche, Tsoar & Kadmon, 2006). All calculations were made with the

221   statistical program R 3.5 (R Core Team, 2017).

222   *Statistical Analysis.* We performed a Mann-Whitney-Wilcoxon test to evaluate differences

223   between forecast and hindcast transferences (Pohlert, 2016). Also, we implemented a Kruskal-

224   Wallis test to evaluate differences in the transference performance between algorithms, when a

225   difference was found, we carried out a Nemenyi test, which makes pairwise multiple

226   comparisons of mean sums of independent samples. These test are implemented in the

227   PMCMR package (Pohlert, 2016). All statistical analyses were performed in R.3.5 (R Core

228   Team, 2017).

229

230   **Results**

231

232   *Species' occurrences.* We found occurrence data for the 117 rodent species endemic to Mexico

233   (Ramírez-Pulido et al., 2014); however, we found sufficient unique records (at least 10) for

234   generating models for each time period only for 44 species. There were no sufficient data for 14

235   species for one of the two periods, whereas there were 59 species for which we did not find 10

236   unique records for both periods. For the 44 species with sufficient data for further analyses,

237   *Peromyscus melanophrys* (n = 504), *Peromyscus difficilis* (n = 440), *Chaetodipus arenarius* (n =

238   248), and *Sigmodon mascotensis* (n=191) held the highest number of records.

239   *Temporal consistency of occurrence data.* We evaluated the consistency of occurrence

240   data at each period to perform temporal transferences for all 44 species with sufficient data. We

241   found that the number of species with high model similarity (>70%) between the two time

242   periods was different depending on the algorithm (Table 1). Maxent got the highest number of

243   species, whereas GAM the lowest (Table 1).

244   *Niche models and model transferences.* In general, all algorithms produced reliable

245   models for calibration in both time periods. Binomial test for models calibrated in the two climatic

246   scenarios were significantly different from random expectations for all species, except for GLM

247   models for *Callospermophilis madrensis*, *Dasyprocta mexicana* and *Dipodomys phillipsii* for T1

248   (1959-1979), and *Dipodomys phillipsii, Neotamias durangae*, *Noetoma goldmani* and *Oryzomys*

249   *guerrerensis* for T2 (1979-2009); a BRT model for *Pappogeomys bulleri* for T2; and a Maxent

250   model for *Dasyprocta mexicana* for T2 (Supplementary Material).

251   We observed differences in algorithm performance when models were transferred

252   across climatic scenarios (fig. 2). Some algorithms performed differently for forecasting and

253   hindcasting: our results showed that Bioclim, GAM and Maxent generally produced higher TSS

254   values for hindcasting than for forecasting, with significant differences between directions only

255   for Maxent ($w = 139$, $p = 0.04$). Similar results were observed for the omission index (OI), with

256   significant differences between directions for Maxent ($w = 140$, $p = 0.043$).

257   We found significant differences between algorithms for transferring models from the

258   present to the past scenario ($H = 17.71$, $p = 0.006$) and from the past to the present ($H = 21.50$,

259    $p < 0.01$), being BRT the algorithm with the highest mean TSS value ($\bar{x} = 0.84 \pm 0.08$) and

260   Bioclim the lowest ($\bar{x} = 0.46 \pm 0.24$) in forecasting. For hindcasting, Maxent obtained the highest

261   mean value ($\bar{x} = 0.82 \pm 0.11$) and Bioclim the lowest ($\bar{x} = 0.60 \pm 0.15$). We observed a similar

262   pattern for OI with significant differences between algorithms for hindcasting ($H = 27.27$, $p =$

263   0.014) and forecasting ($H = 15.93$, $p < 0.01$). We observed the greatest variation in forecasting

264   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

265   consistent behavior and little variation in both directions (forecasting: $\bar{x} = 0.81 \pm 0.09$;

266   hindcasting: $\bar{x} = 0.80 \pm 0.075$), and GAM for hindcasting ($\bar{x} = 0.8 \pm 0.07$).

267   The false negative rate (FNR), which indicates model overestimation, showed

268   significant differences between algorithms for hindcasting ($H = 15.93$, $p = 0.014$) and forecasting

269   ($H = 27.27$, $p < 0.01$). Bioclim showed a difference with BRT ($p = <0.01$) and RF ($p = 0.03$) for

270   forecasting and Maxent ($p = <0.01$) for hindcasting. GAM showed the highest FNR for

271   forecasting ($\bar{x} = 0.29 \pm 0.39$) and GLM for hindcasting ($\bar{x} = 0.10 \pm 0.06$), both algorithms with the

272   highest variation in each direction. In turn, the false positive rate (FPR; an indication of

273   overfitting) was not different for hindcasting (H = 10.96, p = 0.089), but it was so for forecasting

274   (H = 13.22, p=0.039). ~~We observed a difference in~~ FPR between directions only for Maxent (w =

275   301, df = 6, p = 0.04). Conversely, GAM showed the highest FPR values for forecasting.

276

277   **Discussion**

279   ~~Transferences~~ of ecological niche models over time have been widely used to

280   assess the response of biodiversity to climatic changes (Allouche, Tsoar & Kadmon, 2006;

281   Araújo, Thuiller & Pearson, 2006; Rebelo, Tarroso & Jones, 2010; Peterson et al., 2015a).

282   Similar to our results, the variation between algorithms in niche model transfers to alternative

283   climatic scenarios seems to be the rule and has been observed in multiple comparative studies

284   (Moreno-Amat et al., 2015; Beaumont et al., 2016; Bell & Schlaepfer, 2016; García-Callejas &

285   Araújo, 2016; Liang et al., 2018; Qiao et al., 2019). However, to our knowledge, no previous

286   studies have analyzed the effect of occurrence data structure and direction of transfers in the

287   performance of algorithms.

288   In this study, we evaluated the performance and consistency of seven popular

289   algorithms for transferring niche models to past (hindcast) and future (from the past to the

290   present; forecast) climatic conditions for Mexican endemic rodents as study model. We decided

291   to use endemic species for two reasons: (1) the ecological niche of the species is better

292   reconstructed for species with restricted ecological conditions than for those that occupy a great

293   diversity of environments (Raxworthy et al., 2007); and (2) the climatic databases used in this

294   analysis were developed specifically for Mexico (Cuervo-Robayo et al., 2020). Our analyses

295   showed three main results: first, the quality of input data and the algorithm had a significant

296   effect on model output and performance. Second, algorithms did not perform equally for

297   hindcasting and forecasting, where we found generally higher performance for the latter. And

298   third, we found that algorithms performed differently for model transferences, being Random

299   Forest (RF), Boosted Regression Trees (BTR) and Maxent consistently robust, whereas Bioclim

300   showed the lowest performance.

301   Algorithm performance for model transferences has been analyzed under different

302   approaches, including the use of virtual species (García-Callejas & Araújo, 2016) and real

303   species of different biological groups (Heikkinen, Marmion & Luoto, 2012; Beaumont et al., 2016;

304   Fan et al., 2018), as well as testing different sets of variables (Barbet-Massin & Jetz, 2014;

305   Warren et al., 2014) and climate change projections to the past and future (Dobrowski et al.,

306   2011; Veloz et al., 2012). In this study, we assessed algorithm performance for a single

307   taxonomic group under a retrospective approach and found that occurrence data structure

308   between time periods plays a critical role. We know that an adequate sampling is necessary to

309   make a good characterization of the species' ecological niche in the calibration scenario before

310   transferring a model to different scenarios (Araújo, Pearson & Rahbek, 2005), but this is the first

311   time that disparities in the occurrence data between time periods are evaluated in the context of

312   climate change analyses.

313   One important issue when using correlative modeling under a retrospective

314   approach to evaluate the geographic responses of species to climate change is to decouple the

315   effect of data bias or incompleteness from the signal of geographic change. A logical way to

316   proceed when you have two temporally distinct sets of occurrence data and climatic scenarios

317   (i.e., T1 and T2) is to calibrate a model in T1 with the occurrences and climatic scenario from

318   that time period, transfer the resulting model to T2, and use the occurrences of T2 to validate the

319   transfer, and vice versa. However, this approach may have a problem of inconsistency of the

320   occurrences between the two time periods, ergo, if the quantity or quality of occurrence data

321   from the two periods are different, then the two niche models will result somewhat different and

322   transferences will be affected in an unknown manner. In such case, the geographic shift

323   observed under these conditions is mostly due to the occurrence differences, rather than to the

324   responses of species to climatic changes. To avoid this pitfall, our initial step was to test for data

325   consistency between the two time periods by carrying out a cross-modeling procedure. To do so

326   for the 44 species with sufficient data in the two time periods, we first calibrated a model with the

327   climatic scenario and occurrences form T1, and then, we calibrated a second model with

328   occurrences from T2 under the climatic scenario of T1 and compared the two resulting maps,

329   and vice versa. If the data from the two time periods were consistent, the two maps should result

330   highly similar, otherwise they would be significantly different. Interestingly, we found that the

331   number of species with consistent data was different depending on the algorithm, being Maxent

332   the most robust (n = 21), whereas the General Additive Model (GAM) obtained the lowest

333   number (n = 3) (Table 1). This result corroborates the fact that algorithms have different

334   sensitivity to data bias or incompleteness (Moreno-Amat et al., 2015); therefore, if the user

335   knows or suspects about this problem in the dataset, a safer decision is to use a low data-

336   demanding algorithm (e.g., Maxent, Bioclim).

337   There are at least three assumptions when transferring niche models over time: (1)

338   the species is in equilibrium with the environment (Araújo, Pearson & Rahbek, 2005), (2) the

339   species maintains the same response to environmental conditions between time periods

340   (Anderson 2011), and (3) the ecological niche remains conserved along time periods (Soberón &

341   Nakamura, 2009; Wiens et al., 2009; Sequeira et al., 2018). However, other factors have a

342   strong effect on the capacity of niche models to transfer over time, the most critical one being

343   the variability produced by at least three sources: (1) data completeness and structure (i.e.,

344   quality and quantity) (García-Callejas & Araújo, 2016; García- et al., 2023); (2) algorithm

345   performance (Heikkinen, Marmion & Luoto, 2012; Beaumont et al., 2016; Liang et al., 2018;

346   Qiao et al., 2019); and (3) taxonomic idiosyncrasies (Sequeira et al., 2018; Regos et al., 2019).

347   Of the seven algorithms that we tested, Random Forest (RF) showed the greatest predictive

348   capacity and lowest overprediction. This algorithm has proved robust for transferences using

349   virtual species (García-Callejas & Araújo, 2016), good performance for interpolations (Liang et

350   al., 2018), relatively low overprediction (Mi et al., 2017), and high performance (Bell &

351   Schlaepfer, 2016). Our results are consistent with other studies that found BRT having a good

352   performance (García-Callejas & Araújo, 2016; Liang et al., 2018; Sequeira et al., 2018) and high

353   variability (Bell & Schlaepfer, 2016). Maxent was also an algorithm that showed a high

354   performance, particularly for hindcasting. Maxent is the most widely used algorithm

355 (Radosavljevic & Anderson, 2014; Moreno-Amat et al., 2015) in part because it is robust to data

356 biases (Heikkinen, Marmion & Luoto, 2012; Liang et al., 2018; Sequeira et al., 2018). In turn,

357 Bioclim presented the lowest performance, with low overfitting but high overprediction and

358 variability between projections. Bioclim is sensitive to outliers or insufficient sampling, which

359 might be the reason for its poor performance (Dobrowski et al., 2011; Beaumont et al., 2016).

360   Another important and novel result derived from our analyses is that most algorithms

361 performed differently for transferring models to future scenarios than to the past. Bioclim and

362 GAM performed better for hindcasting than for forecasting, whereas RF, Support Vector

363 Machine (SVM) and BRT behaved the opposite. Maxent and the Generalized Linear Model

364 (GLM) were consistent for the two time periods. These differences may reflect the capacity of

365 models to deal with non-analog climates (Pearson et al., 2006). For instance, Bioclim is sensitive

366 to extreme values, so it may be more affected by larger differences between climatic scenarios,

367 as compared to Maxent that has the capacity to clamp and extrapolate into non-analog values

368 (Phillips & Dudík, 2008). More importantly, these results suggests that the choice of an algorithm

369 has implications beyond its simple capacity to produce adequate models in the calibration

370 scenario.

371   Finally, a potential limitation of this study is that the taxonomy of Mexican rodentsis

372 continuously updated. While our research relied on the latest revision of Mexican mammals

373 (Ramírez-Pulido et al. 2014),recent proposals suggest species separation for *Peromyscus*

374 *melanophrys*, *P. furvus*, *P. levipes*, *P. zarhynchus,* and *Osgoogomys banderanus* (Lorenzo et

375 al., 2016; Almendra et al., 2018; Cruz-Gómez et al., 2021; Bradley et al., 2022). As a result, the

376 quantity and spatial structure of occurrences for these species may undergo some modifications.

377 Nonetheless, based on our calibration tests, we believe that the generality of our findings

378 remains unaffected by these taxonomic changes

379   Finally, when the aim of a study is to explore the potential distributional responses of

380 species to climate change under a niche modeling approach, we recommend the following:

381 First, perform a thorough screening and cleaning of occurrence data and identify the most

382   important variables to characterize the ecological niche of the focal species. Then, perform a

383   cross-modeling procedure between time periods in a retrospective fashion (past to present and

384   vice versa) using different algorithms to identify the effect of sampling bias or incompleteness on

385   model transferences, and decide whether the target species is suitable for the analysis. Split the

386   occurrence data into two time periods (e.g. present-near, past) and model between times with

387   different algorithms or perform a consensus to identify the model with the highest transferability

388   over time according to the specific data structure. Finally, identify FNR and FPR to make a

389   decision on the error that best fits the question to be answered by the transfer.

390

Conclusions

391

392   We found that both the quantity and quality of data, along with the choice of the modeling

393   algorithm significantly influence model transferability. Furthermore, we observed that algorithms

394   exhibit unequal performance when transferring models from the past to the present compared to

395   the reverse process. Consequently, we assert that not all species are suitable candidates for

396   transferring models to alternative climatic scenarios; only those species with sufficient

397   occurrence records to produce robust niche models, should be considered for projections time

398   periods.

399   Additionally, not all algorithms perform uniformly across different occurrence data

400   structure and levels of completeness. Some algorithms are more sensitive to small sample sizes

401   and non-analog climates. Therefore, we recommend conducting preliminary tests on algorithmic

402   performance to specific occurrence datasets in a retrospective manner before proceeding with

403   model transferences to future climatic scenarios.

404

Acknowledgements

413

414

References

416   Allouche O, Tsoar A, Kadmon R. 2006. Assessing the accuracy of species distribution models:
417   prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43:1223–
418   1232. DOI: 10.1111/j.1365-2664.2006.01214.x.
419   Almendra AL, González-Cózatl FX, Engstrom MD, Rogers DS. 2018. Evolutionary relationships
420   and climatic niche evolution in the genus *Handleyomys* (Sigmodontinae: Oryzomyini).
421   *Molecular Phylogenetics and Evolution* 128:12–25. DOI: 10.1016/j.ympev.2018.06.018.
422   Anderson RP, Lew D, Peterson AT. 2003. Evaluating predictive models of species' distributions:
423   criteria for selecting optimal models. *Ecological Modelling* 162:211–232. DOI:
424   10.1016/S0304-3800(02)00349-6.
425   Araújo M, Pearson R, Rahbek C. 2005. Equilibrium of species' distribution with climate.
426   *Ecography* 28:693–695.
427   Araújo MB, Peterson T. 2012. Uses and misuses of bioclimatic envelope modeling. *Ecology*
428   93:1527–1539.
429   Araújo M, Thuiller W, Pearson RG. 2006. Climate warming and the decline of amphibians and
430   reptiles in Europe. *Journal of Biogeography* 33:1712–1728. DOI: 10.1111/j.1365-
431   2699.2006.01482.x.
432   Araújo MB, Whittaker RJ, Ladle RJ, Erhard M. 2005. Reducing uncertainty in projections of
433   extinction risk from climate change. *Global Ecology and Biogeography* 14:529–538. DOI:
434   10.1111/j.1466-822X.2005.00182.x.
435   Austin M. 2007. Species distribution models and ecological theory: A critical assessment and
436   some possible new approaches. *Ecological Modelling* 200:1–19. DOI:
437   10.1016/j.ecolmodel.2006.07.005.
438   Barbet-Massin M, Jetz W. 2014. A 40-year, continent-wide, multispecies assessment of relevant
439   climate predictors for species distribution modelling. *Diversity and Distributions*:1–11. DOI:
440   10.1111/ddi.12229.
441   Barbet-Massin M, Jiguet F, Albert CH, Thuiller W. 2012. Selecting pseudo-absences for species
442   distribution models: how, where and how many? *Methods in Ecology and Evolution* 3:327–
443   338. DOI: 10.1111/j.2041-210X.2011.00172.x.
444   Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J,
445   Villalobos F. 2011. The crucial role of the accessible area in ecological niche modeling and

446   species distribution modeling. *Ecological Modelling* 222:1810–1819. DOI:
447   10.1016/j.ecolmodel.2011.02.011.
448 Beaumont LJ, Gallagher R V, Thuiller W, Downey PO, Leishman MR, Hughes L. 2009. Different
449   climatic envelopes among invasive populations may lead to underestimations of current
450   and future biological invasions. :409–420. DOI: 10.1111/j.1472-4642.2008.00547.x.
451 Beaumont LJ, Graham E, Duursma DE, Wilson PD, Cabrelli A, Baumgartner JB, Hallgren W,
452   Esperón-Rodríguez M, Nipperess DA, Warren DL, Laffan SW, VanDerWal J. 2016. Which
453   species distribution models are more (or less) likely to project broad-scale, climate-induced
454   shifts in species ranges? *Ecological Modelling* 342:135–146. DOI:
455   10.1016/j.ecolmodel.2016.10.004.
456 Beaumont LJ, Hughes L. 2002. Potential changes in the distributions of latitudinally restricted
457   Australian butterfly species in response to climate change. *Global Change Biology* 8:954–
458   971. DOI: 10.1046/j.1365-2486.2002.00490.x.
459 Beaumont LJ, Hughes L, Poulsen M. 2005. Predicting species distributions: Use of climatic
460   parameters in BIOCLIM and its impact on predictions of species' current and future
461   distributions. *Ecological Modelling* 186:250–269. DOI: 10.1016/j.ecolmodel.2005.01.030.
462 Bell DM, Schlaepfer DR. 2016. On the dangers of model complexity without ecological
463   justification in species distribution modeling. *Ecological Modelling* 330:50–59. DOI:
464   10.1016/j.ecolmodel.2016.03.012.
465 Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F. 2012. Impacts of climate
466   change on the future of biodiversity. *Ecology letters*:365–377. DOI: 10.1111/j.1461-
467   0248.2011.01736.x.
468 Booth TH, Nix H a., Busby JR, Hutchinson MF. 2014. Bioclim: The first species distribution
469   modelling package, its early applications and relevance to most current MaxEnt studies.
470   *Diversity and Distributions* 20:1–9. DOI: 10.1111/ddi.12144.
471 Bradley RD, Ordóñez-Garza N, Thompson CW, Wright EA, Ceballos G, Kilpatrick CW, Schmidly
472   DJ. 2022. Two new species of *Peromyscus* (Cricetidae: Neotominae) from the Transverse
473   Volcanic Belt of Mexico. *Journal of Mammalogy* 103:255–274. DOI:
474   10.1093/jmammal/gyab128.
475 Breiman L. 2001. Random forests. *Machine Learning* 45:5–32. DOI: 10.1007/978-3-662-56776-
476   0_10.
477 Brotons L, Thuiller W, Araújo MB, Hirzel AH. 2004. Presence-absence versus presence-only
478   modelling methods for predicting bird habitat suitability. *Ecography* 27:437–448.
479 Di Cola V, Broennimann O, Petitpierre B, Breiner FT, D'Amen M, Randin C, Engler R, Pottier J,
480   Pio D, Dubuis A, Pellissier L, Mateo RG, Hordijk W, Salamin N, Guisan A. 2017. ecospat:
481   an R package to support spatial analyses and modeling of species niches and distributions.
482   *Ecography* 40:774–787. DOI: 10.1111/ecog.02671.
483 Cruz-Gómez A, Castro-Campillo A, Ávila-Valle ZA, León-Paniagua L, Ramírez-Sánchez M,
484   Ramírez-Pulido J. 2021. Rejection of the monotypic status of *Peromyscus furvus* (Rodentia:
485   Cricetidae), with consequences for its species group. *Therya* 12:347–367. DOI:
486   10.12933/therya-21-1122.
487 Cuervo-Robayo Á, Ureta C, Gómez-Albores MA, Meneses Mosquera AK, Téllez Valdés O,
488   Martínez-Meyer E. 2020. One hundred years of climate change in Mexico. *PLoS ONE*
489   15:1–19. DOI: 10.1074/jbc.RA118.003614.

490 Dobrowski SZ, Thorne JH, Greenberg JA, Safford HD, Mynsberge AR, Crimmins SM, Swanson
491 AK. 2011. Modeling plant ranges over 75 years of climate change in California, USA:
492 Temporal transferability and species traits. *Ecological Monographs* 81:241–257. DOI:
493 10.1890/10-1325.1.
494 Drake JM, Randin C, Guisan A. 2006. Modelling ecological niches with support vector machines.
495 *Journal of Applied Ecology* 43:424–432. DOI: 10.1111/j.1365-2664.2006.01141.x.
496 Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F,
497 Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura
498 M, Nakazawa Y, Overton JMC, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira
499 R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE. 2006. Novel methods
500 improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151.
501 Elith J, Leathwick JR, Hastie T. 2008. A working guide to boosted regression trees. *Journal of
502 Animal Ecology* 77:802–813. DOI: 10.1111/j.1365-2656.2008.01390.x.
503 Essl F, García-Rodríguez A, Lenzner B, M. A, Capinha C, Gaüzère P, Guisan A, Kühn I, Lenoir
504 J, Richardson DM, Rumpf SB, Svenning J-C, Thuiller W, Zurell D, Dullinger S. 2023.
505 Potential sources of time lags in calibrating species distribution models. *Journal of
506 Biogeography*:1–14. DOI: 10.1111/jbi.14726.
507 Fan JY, Zhao NX, Li M, Gao WF, Wang ML, Zhu GP. 2018. What are the best predictors for
508 invasive potential of weeds? Transferability evaluations of model predictions based on
509 diverse environmental data sets for Flaveria bidentis. *Weed Research* 58:141–149. DOI:
510 10.1111/wre.12292.
511 Fielding AH, Bell JF. 1997. A review of methods for the assessment of prediction errors in
512 conservation presence/absence models. *Environmental Conservation* 24:38–49. DOI:
513 10.1017/S0376892997000088.
514 García-Callejas D, Araújo MB. 2016. The effects of model and data complexity on predictions
515 from species distributions models. *Ecological Modelling* 326:4–12. DOI:
516 10.1016/j.ecolmodel.2015.06.002.
517 Garcia RA, Cabeza M, Altwegg R, Araújo MB. 2016. Do projections from bioclimatic envelope
518 models and climate change metrics match? *Global Ecology and Biogeography* 25:65–74.
519 DOI: 10.1111/geb.12386.
520 Golicher D, Ford A, Cayuela L, Newton A. 2012. Pseudo-absences, pseudo-models and
521 pseudo-niches: pitfalls of model selection based on the area under the curve. *International
522 Journal of Geographical Information Science* 26:2049–2063. DOI:
523 10.1080/13658816.2012.719626.
524 Guisan A, Edwards TC, Hastie T. 2002. Generalized linear and generalized additive models in
525 studies of species distributions: setting the scene. *Ecological Modelling* 157:89–100. DOI:
526 10.1016/S0304-3800(02)00204-1.
527 Habibullah MS, Din BH, Tan SH, Zahid H. 2022. Impact of climate change on biodiversity loss:
528 global evidence. *Environmental Science and Pollution Research* 29:1073–1086. DOI:
529 10.1007/s11356-021-15702-8.
530 Heikkinen RK, Marmion M, Luoto M. 2012. Does the interpolation accuracy of species
531 distribution models come at the expense of transferability? *Ecography* 35:276–288. DOI:
532 10.1111/j.1600-0587.2011.06999.x.

533   Hernandez PA, Graham CH, Master LL, Albert DL. 2006. The effect of sample size and species
534   characteristics on performance of different species distribution modeling methods.
535   *Ecography* 29:773–785. DOI: 10.1111/j.0906-7590.2006.04700.x.
536   Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated
537   climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978.
538   DOI: 10.1002/joc.1276.
539   Hijmans RJ, Graham C. 2006. The ability of climate envelope models to predict the effect of
540   climate change on species distributions. *Global Change Biology* 12:2272–2281. DOI:
541   10.1111/j.1365-2486.2006.01256.x.
542   Hijmans ARJ, Phillips S, Leathwick J, Elith J, Hijmans MRJ. 2017. Package ' dismo .' *CRAN*.
543   Available at https://cran.r-project.org/web/packages/dismo/index.html (accessed 18 may
544   2022).
545   Jaroszynska F, Rixen C, Woodin S, Lenoir J, Wipf S. 2023. Resampling alpine herbarium
546   records reveals changes in plant traits over space and time. *Journal of Ecology* 111:338–
547   355. DOI: 10.1111/1365-2745.14062.
548   Jiménez-Valverde A, Lobo JM, Hortal J. 2008. Not as good as they seem: the importance of
549   concepts in species distribution modelling. *Diversity and Distributions* 14:885–890. DOI:
550   10.1111/j.1472-4642.2008.00496.x.
551   Kharouba HM, Algar AC, Kerr JT. 2009. Historically calibrated predictions of butterfly species'
552   range shift using global change as a pseudo-experiment. *Ecology* 90:2213–2222. DOI:
553   10.1890/08-1304.1.
554   Liang W, Papeş M, Tran L, Grant J, Washington-Allen R, Stewart S, Wiggins G. 2018. The effect
555   of pseudo-absence selection method on transferability of species distribution models in the
556   context of non-adaptive niche shift. *Ecological Modelling* 388:1–9. DOI:
557   10.1016/j.ecolmodel.2018.09.018.
558   Lorenzo C, Álvarez-Castañeda ST, Pérez-Consuegra SG, Patton JL. 2016. Revision of the
559   Chiapan deer mouse, *Peromyscus zarhynchus*, with the description of a new species.
560   *Journal of Mammalogy* 97:910–918. DOI: 10.1093/jmammal/gyw018.
561   MacLean SA, Rios Dominguez AF, de Valpine P, Beissinger SR. 2018. A century of climate and
562   land-use change cause species turnover without loss of beta diversity in California's Central
563   Valley. *Global Change Biology* 24:5882–5894. DOI: 10.1111/gcb.14458.
564   Maiorano L, Falcucci A, Zimmermann NE, Psomas A, Pottier J, Baisero D, Rondinini C, Guisan
565   A, Boitani L. 2011. The future of terrestrial mammals in the Mediterranean basin under
566   climate change. *Philosophical Transactions of the Royal Society B: Biological Sciences*
567   366:2681–2692. DOI: 10.1098/rstb.2011.0121.
568   Martínez-Meyer E, Peterson AT, Hargrove WW. 2004. Ecological niches as stable distributional
569   constraints on mammal species, with implications for Pleistocene extinctions and climate
570   change projections for biodiversity. *Global Ecology and Biogeography* 13:305–314.
571   Merow C, Smith MJ, Edwards Jr TC, Guisan A, McMahon SM, Normand S, Thuiller W, Wüest
572   RO, Zimmermann NE, Elith J. 2014. What do we gain from simplicity versus complexity in
573   species distribution models? *Ecography* 37:1–15. DOI: 10.1111/ecog.00845.
574   Merow C, Smith MJ, Silander J a. 2013. A practical guide to MaxEnt for modeling species'
575   distributions: what it does, and why inputs and settings matter. *Ecography* 36:1058–1069.
576   DOI: 10.1111/j.1600-0587.2013.07872.x.

577   Mi C, Huettmann F, Guo Y, Han X, Wen L. 2017. Why choose Random Forest to predict rare
578   species distribution with few samples in large undersampled areas? Three Asian crane
579   species models provide supporting evidence. *PeerJ*. DOI: 10.7717/peerj.2849.
580   Moreno-Amat E, Mateo RG, Nieto-Lugilde D, Morueta-Holme N, Svenning JC, García-Amorena
581   I. 2015. Impact of model complexity on cross-temporal transferability in Maxent species
582   distribution models: An assessment using paleobotanical data. *Ecological Modelling*
583   312:308–317. DOI: 10.1016/j.ecolmodel.2015.05.035.
584   Moritz C, Patton JL, Conroy CJ, Parra JL, White GC, Beissinger SR. 2008. Impact of a century
585   of climate change on small mammal communities in Yosemite national park, USA. *Science*
586   322:261–264. DOI: 10.1126/science.1163428.
587   Muscarella R, Galante PJ, Soley-Guardia M, Boria RA, Kass JM, Uriarte M, Anderson RP. 2014.
588   ENMeval: An R package for conducting spatially independent evaluations and estimating
589   optimal model complexity for Maxent ecological niche models. *Methods in Ecology and*
590   *Evolution* 5:1198–1205. DOI: 10.1111/2041-210X.12261.
591   Naimi B, Araújo MB. 2016. SDM: A reproducible and extensible R platform for species
592   distribution modelling. *Ecography* 39:368–375. DOI: 10.1111/ecog.01881.
593   Nix HA. 1986. *A biogeographic analysis of Australian Elapid snakes. In: Longmore, R.* Canberra.
594   Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC,
595   D'amico J a., Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y,
596   Lamoreux JF, Wettengel WW, Hedao P, Kassem KR. 2001. Terrestrial Ecoregions of the
597   World: A New Map of Life on Earth. *BioScience* 51:933. DOI: 10.1641/0006-
598   3568(2001)051[0933:TEOTWA]2.0.CO;2.
599   Pacifici M, Foden WB, Visconti P, Watson JEM, Butchart SHM, Kovacs KM, Scheffers BR, Hole
600   DG, Martin TG, Akçakaya HR, Corlett RT, Huntley B, Bickford D, Carr JA, Hoffmann AA,
601   Midgley GF, Pearce-Kelly P, Pearson RG, Williams SE, Willis SG, Young B, Rondinini C.
602   2015. Assessing species vulnerability to climate change. *Nature Climate Change* 5:215–
603   225. DOI: 10.1038/nclimate2448.
604   Parmesan C, Yohe G. 2003. A globally coherent fingerprint of climate change impacts across
605   natural systems. *Nature* 421:37–42. DOI: 10.1038/nature01286.
606   Pearson RG. 2006. Climate change and the migration capacity of species. *Trends in Ecology*
607   *and Evolution* 21:111–113. DOI: 10.1016/j.tree.2005.11.022.
608   Pearson RG, Dawson TP. 2003. Predicting the impacts of climate change on the distribution of
609   species: are bioclimate envelope models useful? *Global Ecology and Biogeography*
610   12:361–371. DOI: 10.1046/j.1466-822X.2003.00042.x.
611   Peterson AT. 2011. Ecological niche conservatism: A time-structured review of evidence.
612   *Journal of Biogeography* 38:817–827. DOI: 10.1111/j.1365-2699.2010.02456.x.
613   Peterson AT, Martínez-Meyer E, González-Salazar C. 2004. Reconstructing the Pleistocene
614   geography of the Aphelocoma jays (Corvidae). *Diversity and Distributions* 10:237–246.
615   DOI: https://doi.org/10.1111/j.1366-9516.2004.00097.x.
616   Peterson AT, Navarro-Siguenza AG, Martinez-Meyer E, Cuervo-Robayo AP, Berlanga H,
617   Soberón J. 2015a. Twentieth century turnover of Mexican endemic avifaunas: Landscape
618   change versus climate drivers. *Science Advances* 1:e1400071–e1400071. DOI:
619   10.1126/sciadv.1400071.

620 Peterson AT, Navarro-Siguenza AG, Martínez-Meyer E, Cuervo-Robayo AP, Berlanga H,
621   Soberón J. 2015b. Twentieth century turnover of Mexican endemic avifaunas: Landscape
622   change versus climate drivers. *Science Advances* 1:e1400071–e1400071. DOI:
623   10.1126/sciadv.1400071.
624 Phillips SJ. 2008. Transferability, sample selection bias and background data in presence-only
625   modelling: a response to Peterson et al. (2007). *Ecography*:080227084236895. DOI:
626   10.1111/j.2007.0906-7590.05378.x.
627 Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species
628   geographic distributions. *Ecological Modelling* 190:231–259. DOI:
629   10.1016/j.ecolmodel.2005.03.026.
630 Piirainen S, Lehikoinen A, Husby M, Kålås JA, Lindström Å, Ovaskainen O. 2023. Species
631   distributions models may predict accurately future distributions but poorly how distributions
632   change: A critical perspective on model validation. *Diversity and Distributions* 29:654–665.
633   DOI: 10.1111/ddi.13687.
634 Pohlert T. 2016. The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR).
635   *CRAN*:1–27.
636 Prasad AM, Iverson LR, Liaw A. 2006. Newer Classification and Regression Tree Techniques :
637   Bagging and Random Forests for Ecological Prediction. *Ecosystems* 9:181–199. DOI:
638   10.1007/s10021-005-0054-1.
639 Qiao H, Feng X, Escobar LE, Peterson AT, Soberón J, Zhu G, Papeş M. 2019. An evaluation of
640   transferability of ecological niche models. *Ecography* 42:521–534. DOI:
641   10.1111/ecog.03986.
642 R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. DOI:
643   10.1007/978-3-540-74686-7.
644 Radosavljevic A, Anderson RP. 2014. Making better Maxent models of species distributions:
645   Complexity, overfitting and evaluation. *Journal of Biogeography* 41:629–643. DOI: 10.1111/
646   jbi.12227.
647 Ramírez-Pulido J, González-Ruiz N, Gardner AL, Arroyo-Cabrales J. 2014. *List of recent land
648   mammals of Mexcio*. Lubbock, Texas. DOI: 10.1177/002194368302000404.
649 Raxworthy CJ, Ingram CM, Rabibisoa N, Pearson RG. 2007. Applications of ecological niche
650   modeling for species delimitation: A review and empirical evaluation using day geckos
651   (Phelsuma) from Madagascar. *Systematic Biology* 56:907–923. DOI:
652   10.1080/10635150701775111.
653 Rebelo H, Tarroso P, Jones G. 2010. Predicted impact of climate change on European bats in
654   relation to their biogeographic patterns. *Global Change Biology* 16:561–576. DOI:
655   10.1111/j.1365-2486.2009.02021.x.
656 Regos A, Gagne L, Alcaraz-Segura D, Honrado JP, Domínguez J. 2019. Effects of species traits
657   and environmental predictors on performance and transferability of ecological niche
658   models. *Scientific Reports* 9:1–14. DOI: 10.1038/s41598-019-40766-5.
659 Rubidge EM, Monahan WB, Parra JL, Cameron SE, Brashares JS. 2011. The role of climate,
660   habitat, and species co-occurrence as drivers of change in small mammal distributions over
661   the past century. *Global Change Biology* 17:696–708. DOI: 10.1111/j.1365-
662   2486.2010.02297.x.

663 Sequeira AMM, Bouchet PJ, Yates KL, Mengersen K, Caley MJ. 2018. Transferring biodiversity
664   models for conservation: Opportunities and challenges. *Methods in Ecology and Evolution*
665   9:1250–1264. DOI: 10.1111/2041-210X.12998.
666 Sillero N. 2011. What does ecological modelling model? A proposed classification of ecological
667   niche models based on their underlying methods. *Ecological Modelling* 222:1343–1346.
668   DOI: 10.1016/j.ecolmodel.2011.01.018.
669 Sillero N, Campos JC, Arenas-Castro S, Barbosa AM. 2023. A curated list of R packages for
670   ecological niche modelling. *Ecological Modelling* 476. DOI:
671   10.1016/j.ecolmodel.2022.110242.
672 Soberón J, Nakamura M. 2009. Niches and distributional areas : Concepts, methods and
673   assumptions. *PNAS* 106:19644–19650.
674 Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BFN,
675   De Siqueira MF, Grainger A, Hannah L, Hughes L, Huntley B, Van Jaarsveld AS, Midgley
676   GF, Miles L, Ortega-Huerta M a, Peterson AT, Phillips OL, Williams SE. 2004. Extinction
677   risk from climate change. *Nature* 427:145–8. DOI: 10.1038/nature02121.
678 Vaz UL, Cunha HF, Nabout JC. 2015. Trends and biases in global scientific literature about
679   ecological niche models. *Brazilian Journal of Biology* 75:17–24. DOI: 10.1590/1519-
680   6984.22713.
681 Veloz SD, Williams JW, Blois JL, He F, Otto-Bliesner B, Liu Z. 2012. No-analog climates and
682   shifting realized niches during the late quaternary: Implications for 21st-century predictions
683   by species distribution models. *Global Change Biology* 18:1698–1713. DOI:
684   10.1111/j.1365-2486.2011.02635.x.
685 Waltari E, Guralnick RP. 2009. Ecological niche modelling of montane mammals in the Great
686   Basin, North America: examining past and present connectivity of species across basins
687   and ranges. *Journal of Biogeography* 36:148–161. DOI: 10.1111/j.1365-
688   2699.2008.01959.x.
689 Warren DL, Wright AN, Seifert SN, Shaffer HB. 2014. Incorporating model complexity and
690   spatial sampling bias into ecological niche models of climate change risks faced by 90
691   California vertebrate species of concern. *Diversity and Distributions* 20:334–343. DOI:
692   10.1111/ddi.12160.
693 Warton D, Aarts G. 2013. Advancing our thinking in presence-only and used-available analysis.
694   *Journal of Animal Ecology* 82:1125–1134. DOI: 10.1111/1365-2656.12071.
695 Werkowska W, Márquez AL, Real R, Acevedo P. 2017. A practical overview of transferability in
696   species distribution modeling. *Environmental Reviews* 25:127–133. DOI: 10.1139/er-2016-
697   0045.
698 Widick I V., Bean WT. 2019. Evaluating current and future range limits of an endangered,
699   keystone rodent (*Dipodomys ingens*). *Diversity and Distributions* 25:1074–1087. DOI:
700   10.1111/ddi.12914.
701 Wiens JA, Stralberg D, Jongsomjit D, Howell CA, Snyder MA. 2009. Niches, models, and climate
702   change: assessing the assumptions and uncertainties. *PNAS* 106 Suppl:19729–19736.
703   DOI: 10.1073/pnas.0901639106.
704 Yates KL, Bouchet PJ, Caley MJ, Mengersen K, Randin CF, Parnell S, Fielding AH, Bamford AJ,
705   Ban S, Barbosa AM, Dormann CF, Elith J, Embling CB, Ervin GN, Fisher R, Gould S, Graf
706   RF, Gregr EJ, Halpin PN, Heikkinen RK, Heinänen S, Jones AR, Krishnakumar PK, Lauria

707   V, Lozano-Montes H, Mannocci L, Mellin C, Mesgaran MB, Moreno-Amat E, Mormede S,
708   Novaczek E, Oppel S, Ortuño Crespo G, Peterson AT, Rapacciuolo G, Roberts JJ, Ross
709   RE, Scales KL, Schoeman D, Snelgrove P, Sundblad G, Thuiller W, Torres LG, Verbruggen
710   H, Wang L, Wenger S, Whittingham MJ, Zharikov Y, Zurell D, Sequeira AMM. 2018.
711   Outstanding Challenges in the Transferability of Ecological Models.
712

Table 1(on next page)

Species with sufficient occurrences

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

Algorithm	Number of species with consistent models	
	Occurrences of period 1 modeled with climatic variables of Period 2	Occurrences of period 2 modeled with climatic variables of Period 1
Bioclim	41	42
Boosted Regression Trees (BRT)	44	43
Generalized Additive Models (GAM)	44	44
Generalized Linear Models (GLM)	44	44
Maxent	44	44
Random Forest (RF)	44	44
Support Vector Machine (SVM)	44	44

Figure 1

Methodological approach

Methodological approach used 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.
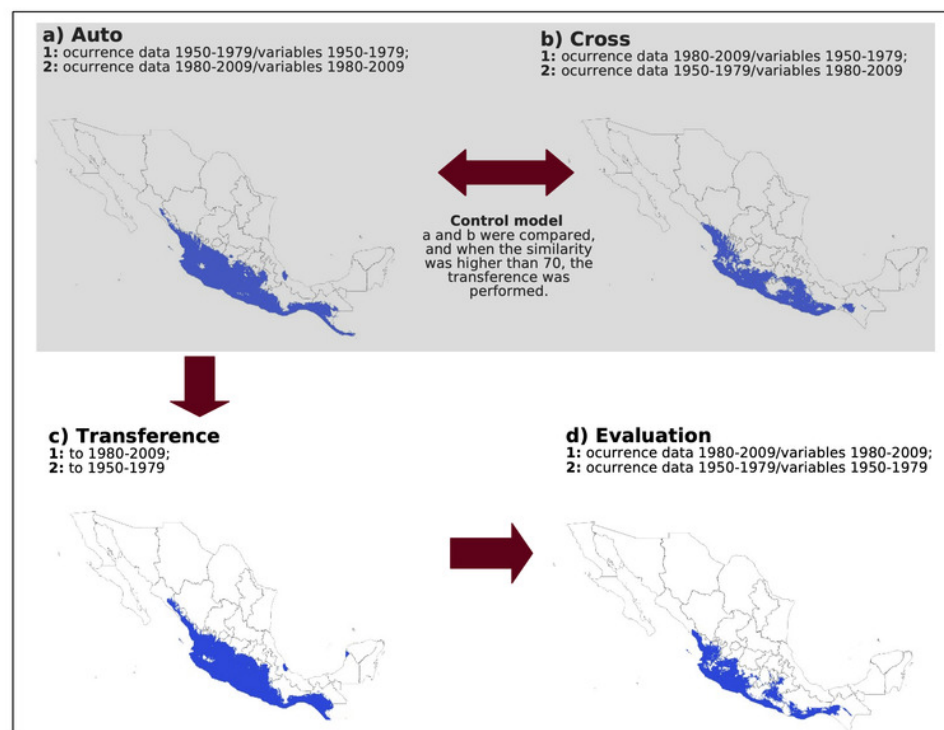
Figure 2

Model performance

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.