



Too much of a good thing? Supplementing current species observations with fossil data to assess climate change vulnerability via ecological niche models

Arianna M. Belfiore^{a,1}, Alessandro Mondanaro^{b,1}, Silvia Castiglione^c, Marina Melchionna^c, Giorgia Girardi^c, Pasquale Raia^c, Mirko Di Febbraro^{a,*}

^a EnviXLab, Department of Biosciences and Territory, University of Molise, Pesche, Italy

^b Department of Earth Sciences, University of Florence, Florence, Italy

^c Department of Earth Sciences, Environment and Resources, University of Naples Federico II, Naples, Italy

ARTICLE INFO

Keywords:

Ecological niche models
Climate change
Biodiversity conservation
Geographic range shift
Fossil record

ABSTRACT

Ecological niche models (ENMs) are a powerful tool in ecological research and conservation planning. Since ENMs provide probability maps of suitable areas under environmental change, they may assist in designing conservation actions and addressing conservation priorities. However, ENMs are usually implemented by learning the species climatic preferences from their current geographic distribution, which leaves them vulnerable to the issue of niche truncation, as it comes with non-climatic limits to the current species distribution as posed by anthropic activities and settlements, and competition avoidance. These problems might be alleviated by the inclusion of fossil occurrences, which refer to moments during species existence when such limits were absent or distributed differently. By stacking species occurrences from different time moments is conceivable that a larger fraction of the species fundamental niche is thence explored. Here, we combined current and fossil occurrence data for 38 medium-large mammal species of conservation concern in order to assess the influence of the fossil record inclusion on ENM predictions under future climate change scenarios. We found that ignoring or including fossil data yields consistent trends in terms of predicted range increase/decrease. Yet, although adding fossil data invariably results in increased niche width, estimates of range change improved for almost a half only of the species pool. These results suggest that most species might currently be in non-equilibrium with their environment, and that the inclusion of fossil data may be crucial to the better understanding of species climatic requirements, hence for designing more effective conservation strategies.

1. Introduction

Ecological niche models (ENMs) are sought to represent the bioclimatic preferences and tolerance limits of species (Muscarella et al., 2014). One key feature of ENMs is the ability to produce probability maps of suitable areas for the species, either under current climatic conditions, in the past, or in the future. This feature makes ENMs a viable tool to model future species response to predicted climate change (Barbet-Massin et al., 2018; Fois et al., 2018; Frans et al., 2022; Hanson et al., 2020). Yet, it must be considered that projecting ENMs in the

future (or in the past) stands on strong assumptions (Elith and Leathwick, 2009; Guisan et al., 2014). First, ENMs assume that the species occurrences used to produce the model capture a faithful representation of the physiological tolerance spectrum of the species. That is, the species is presumed to be well-sampled and at equilibrium with its environment. In areas under intense human impact, such supposition could be severely counterfeited, as anthropic activities and settlements may limit the extent of area available to the species (Chiarenza et al., 2023; Di Marco et al., 2021; Frans et al., 2022; Maiorano et al., 2019; Saupe et al., 2019). Secondly, the reliability of ENMs critically depends on the

* Corresponding author.

E-mail addresses: a.belfiore@studenti.unimol.it (A.M. Belfiore), alessandro.mondanaro@unifi.it (A. Mondanaro), silvia.castiglione@unina.it (S. Castiglione), marina.melchionna@unina.it (M. Melchionna), giorgia.girardi@unina.it (G. Girardi), pasquale.raia@unina.it (P. Raia), mirko.difebbraro@unimol.it (M. Di Febbraro).

¹ These authors contributed equally.

number of occurrence data available (Jiménez-Valverde, 2020; Santini et al., 2021), on the natural temporal variability in climates (Santini et al., 2021), and on model transferability to past or future climates (Qiao et al., 2019). As a matter of fact, ENMs derived from the current distribution of a species represent its realized climatic niche, but ENM projections in time require approaching the species bioclimatic potential, that is its fundamental niche. Obvious workarounds to sort these problems out are using ENM strategies that are little sensitive to limited entry data (Breiner et al., 2015; Lomba et al., 2010), which are not entirely based on the current distribution of the species (Mondanaro et al., 2023; Qiao et al., 2017) and supplementing the current distribution with fossil observations and their related climatic data (Fritz et al., 2013; Jones et al., 2019; Maguire et al., 2015; Maiorano et al., 2013; Raia et al., 2020). The inclusion of fossil data is notoriously important to increase the quality of inference in studies of phenotypic evolution (Castiglione et al., 2020; Slater et al., 2012). Since the climatic niche can be assimilated to an evolving phenotypic trait (Holt and Gaines, 1992; Rolland et al., 2018), it seems reasonable to assume that including fossil observations and their climates could ameliorate ENM predictions, especially for the rare species, by informing the model about the climatic conditions potentially suitable to the species that are not currently sampled, hence approaching the species fundamental climatic niche. So far, the few studies explicitly integrating modern and fossil data into climate change vulnerability assessments via ENMs were consistent in reporting better predicted outcomes from models including fossil data (Lima et al., 2017; Sales et al., 2022). This evidence seems to suggest that modern occurrence data might be representative of a limited fraction of the species climatic niche in most cases, thus yielding overestimated vulnerability to climate change (Chiarenza et al., 2023; Sales et al., 2022).

That said, the benefits derived from coupling modern and fossil data into ENMs to predict climate change effects are not unambiguous (Moreno-Amat et al., 2017). First, if the bioclimatic niche evolved, the fossil data may represent climates that were, but no longer are, suitable to the species. Still, fossil data might be representative of a capacity to explore a fraction of the natural variability in climates that is no longer existent. For instance, lions *Panthera leo* were once as widespread as to occur in icy North America during the last glacial maximum (Barnett et al., 2009). However, this is no proof that extant lion individuals may survive in periglacial areas, and the species passed through severe bottlenecks that may have reduced its bioclimatic tolerance limits by depleting its gene pool (Barnett et al., 2009).

Herein, we aim at exploring whether and how the inclusion of fossil data into climate change vulnerability assessments performed via ENMs might alter predicted effects on species distribution, in comparison to models trained exclusively with modern occurrence data. We started from two related hypotheses: i) adding fossil data to modern occurrences provides a steady increase in the sampled climatic niche, that widens the species calculated climatic tolerance; ii) that these wider climatic niches bring about reduced negative effects by climate change on the future geographic distribution in most species. We tested both hypotheses by focusing on a number of terrestrial mammal species of conservation concern, currently listed as “critically endangered”, “endangered”, “vulnerable” or “near threatened” by the IUCN. The two study hypotheses were tested pursuing five specific objectives: i) calculate niche width and overlap between modern and fossil data, identifying the major climate conditions characterizing modern and fossil realized niches; ii) calibrate two groups of ENMs, relying on modern occurrence only (i.e., modern ENMs) and modern plus fossil data pooled together (i.e., full ENMs), projecting both groups to year 2080 under two different climate change scenarios; iii) identify possible species showing discrepancies in current–2080 range modifications between modern and full ENMs predictions; iv) test for the relationship between range net change values obtained by modern and full ENMs and the niche width gain generated by adding fossil data to modern occurrences; and v) evaluate whether possible discrepancies in predicted range alterations

between the two ENM groups depend on different past climates being sampled.

2. Methods

2.1. Modern and fossil occurrences

We collected modern occurrence data from two online databases, the “Global Biodiversity Information Facility” (GBIF; www.gbif.org/; Table S1) and iNaturalist (<https://www.inaturalist.org/>). The accuracy of records gathered from GBIF was checked by including only data provided with at least two decimal places (0.01 decimal degrees, corresponding to 1.11 km at the equator), while data from iNaturalist were assessed by removing records without photos attached and expert confirmation. Data were also filtered by removing duplicates and records with unrealistic coordinates, as well as occurrences from natural history collections (i.e., the so-called “preserved” data, according to the GBIF vocabulary; Marcer et al., 2022). As to fossil data, we implemented the mammal database provided in Mondanaro et al. (2021), where radiocarbon dates were calibrated using the “Bchron” R package (Hasslett and Parnell, 2008) through the “Intcal20” calibration curve (Reimer et al., 2020). From the initial set of candidate mammal species, we excluded those reporting <20 modern occurrence data, retaining a total of 38 species (Table S1). After filtering procedures, we retrieved 15,012 modern occurrence records and 1,834 fossil data. Moreover, fossil data for each species were replicated 10 times as to account for the uncertainty in age estimate inherent to individual fossil data (Mondanaro et al., 2021). Specifically, at each replicate the age of each fossil site was drawn from a uniform distribution ranging from the minimum to the maximum age estimate of the site.

2.2. Climate predictors

As environmental predictors, we considered 17 bioclimatic variables. Specifically, variables for training modern ENMs were gathered from the CHELSA database version 1.2 (Karger et al., 2017), while past climatic predictors were obtained by Krapp et al. (2021). The latter database includes paleoclimate simulations generated by the HadCM3 global circulation model (Valdes et al., 2017), covering the last 800 kya at 1 kya temporal resolution. Since paleoclimate simulations in Krapp et al. (2021) did not model diurnal cycle, both annual mean diurnal range (BIO2) and isothermality (BIO3) were unavailable, thus reducing the number of bioclimatic variables to 17 in lieu of 19 (Karger et al., 2017). For ENMs calibration, CHELSA variables were rasterized at 10 km spatial resolution, while past climatic predictors were obtained at 50 km spatial resolution. For prediction purposes, both current and future climate variables were aggregated at 50 km spatial resolution. Bioclimatic variables were further checked for multicollinearity by calculating the variance inflation factor (VIF) on pooled predictor values over all the 800 kya temporal span, as to detect possible changes in correlation structure over time. After posing a VIF threshold ≤ 5 (Zuur et al., 2007), we retained 6 predictors: Temperature seasonality (BIO4), Mean daily maximum air temperature of the warmest month (BIO5), Mean daily mean air temperatures of the wettest quarter (BIO8), Precipitation amount of the driest month (BIO14), Mean monthly precipitation amount of the wettest quarter (BIO16), Mean monthly precipitation amount of the driest quarter (BIO17).

2.3. Niche overlap analyses

Comparison of the modern and full (i.e., fossil plus modern) climatic niches was carried out using the analytical framework proposed by Broennimann et al. (2012). Accordingly, PCA was used to decompose the environmental space defined for modern and fossil data (i.e., all the environmental conditions intersected by the occurrence and background points). Occurrence records and climate conditions were projected into

this PCA space, then their densities were computed across the first two PCs relying on a kernel density estimator. Densities of occurrence and background environments were then divided by the maximum number of occurrences in any cell of the environmental space and by the number of sites with the most common climate, respectively (Broennimann et al., 2012). The process generated a density grid in the environmental space that was used to quantify niche overlap between modern and fossil niches in terms of Schoener's *D* index (Schoener, 1970). This index ranges from 0 to 1 with values of 0 indicating no niche overlap and values of 1 indicating a complete overlap. For each species, the entire procedure was repeated for each of the 10 replicated fossil datasets generated to account for the uncertainty in age estimate. Niche overlap analyses were carried out using the "ecospat" R package (Di Cola et al., 2017).

2.4. Ecological niche models

For modern ENMs, we generated 10,000 background points for each species in the study area, covering a region identified by all the WWF Terrestrial Ecoregions (Olson et al., 2001). For full ENMs, we generated a single ensemble model for each of the 10 replicated fossil datasets. For each species modelled under full ENMs, any single replicate included modern data, one of the replicated fossil datasets and 10,000 background points selected within the WWF Ecoregions. To compensate for potential sampling biases, background points were geographically located according to the density of the occurrence data, so that more background points occur where presences are denser (Mondanaro et al., 2021; Roy-Dufresne et al., 2019; Syfert et al., 2013). For full ENMs, we divided the record of each species into 1000-years long consecutive time bins and allocated the 10,000 background points proportionally to the number of presences per time bin. For both modern and full ENMs, we adopted an ensemble forecasting approach by relying on the functionalities provided in the 'biomod2' R package (Thuiller et al., 2009). Models were trained by using four algorithms: Generalized Linear Model (GLM), Generalized Boosting Model (GBM), Random Forest (RF) and Maximum Entropy (MAXENT). To evaluate the predictive accuracy of modern ENMs, we primarily adopted a spatial block cross-validation scheme (Roberts et al., 2017), where occurrences falling into three out of four folds were used for model calibration, while the held-out data were used for evaluation. Furthermore, we assessed modern ENMs temporal transferability (Roberts et al., 2017) through a temporal block cross-validation approach, thus dividing the fossil data of each species into 10 time bins of approximately the same length and projecting modern ENMs against each bin. A similar approach was applied to evaluate full ENMs, at this time using nine out of 10 bins for calibration and the held-out ones to assess predictive performance. ENMs predictive accuracy was assessed by calculating the area under the receiver operating characteristic curve (AUC; Swets, 1988) and the continuous Boyce index (CBI; Hirzel et al., 2006). To avoid using poorly calibrated models, we considered in further analyses only predictions derived from models with $AUC \geq 0.7$. Model averaging was performed by weighting the individual model projections by their AUC values and averaging the results (Marmion et al., 2009). Both modern and full ENMs were projected on current climate and two future climate change scenarios (i.e., RCP4.5 and RCP8.5) generated under the HadGEM2-CC global circulation model (Martin et al., 2011; see Beyer and Manica, 2020 for a similar pairing of HadCM3 and HadGEM2 families of global circulation models). The future scenarios predict either mild (RCP4.5) or severe (RCP8.5) climate change in 2080. Current and future ENMs projections were binarized to obtain range maps according to three thresholding approaches (i.e., 'equalize sensitivity and specificity,' 'maximize TSS', and '10th percentile of predicted probability'; Mondanaro et al., 2021), as to account for the effect of using different binarization schemes (Di Febbraro et al., 2019). For each species, both suitability and binary spatially explicit predictions were constrained within the boundaries of the same WWF Ecoregions selected to delimit the area of background

points (Newbold et al., 2015). To prevent model extrapolation effect (i.e., predicting on covariate values lying outside the calibration range), ENMs projections were constrained using environmental clamping (Elith et al., 2011), that is capping covariate values at the training range limits. Climate change effect on species distribution was quantified by calculating the range net change (RNC, hereafter) metric (in terms of percentage of gain/loss suitable territory calculated comparing the current to the future range) on binary maps generated for each species, model and scenario (Franklin et al., 2013). Lastly, we compared range net change values predicted by modern and full ENMs, and classified the species in two groups: "higher fossil species" (i.e., range net change from full ENMs is higher than that from modern ENMs) and "higher modern species" (i.e., showing the opposite figure).

2.5. Range net change versus niche width gain granted by fossil data

To test whether an increase in niche width generated by the inclusion of fossil data is significantly correlated to a reduction of the predicted climate change effects on species distribution, we fit a Linear Mixed Model (LMM). Specifically, we considered the difference between RNC values predicted by full ENMs and modern ENMs (ΔRNC) as the response variable, and the percentage increase in niche width provided by fossil data as explanatory variable, in interaction with the climate change scenario (RCP4.5 vs. RCP8.5) and the species outcome (that is either "higher modern" or "higher fossil"). Furthermore, we added model replicate (i.e., due to fossil age uncertainty) and binarization threshold as random factors, allowing the model to vary its intercept accordingly. We also fit a second LMM dropping the interaction with species outcome, as to quantify the contribution of this specific term on the relationship between the increase in niche width generated by fossil data and ΔRNC . The two alternative LMMs were compared to each other through Akaike Information Criterion (AIC). LMM goodness-of-fit was assessed through conditional R^2 (Nakagawa and Schielzeth, 2013). All the analyses were performed with the "lme4" (Bates et al., 2015), "lmerTest" (Kuznetsova et al., 2017) and "MuMIn" (Bartoń, 2016) R packages.

2.6. Environmental differences in fossil niche between "higher modern" and "higher fossil" species

To test if a diverging pattern in predicted climate change effect between "higher modern" and "higher fossil" species might depend on differences in the climatic conditions provided by the fossil niches of these species, we deployed a Random Forest (RF; Breiman, 2001) classification model. Specifically, we pooled all the species together and set their outcome ("higher modern" or "higher fossil") as response variable and the climate values associated with fossil niche portions in the PCA environmental space as covariates. We calibrated a separate RF model for each of the two climate change scenarios (RCP4.5 and RCP8.5), assessing classification performance as the accuracy rate calculated through a five folds cross-validation approach (Liaw and Wiener, 2002). We also generated marginal response plots as to depict the shape of the relationship between the climatic conditions provided by the fossil niches and the probability for a specie to be classified as "higher fossil" or "higher modern". All the analyses were carried out using the "randomForest" R package (Liaw and Wiener, 2002).

Lastly, we devised a procedure to test whether the inclusion of fossil data generated exceedingly wide climatic niches, which would be more indicative of intraspecific niche evolution rather than of a fuller representation of the species fundamental niche. Once this procedure indicated for each species the most climatically divergent time bins, all the analyses were entirely re-run after excluding all the fossil data older than these selected intervals (Supplementary text S1).

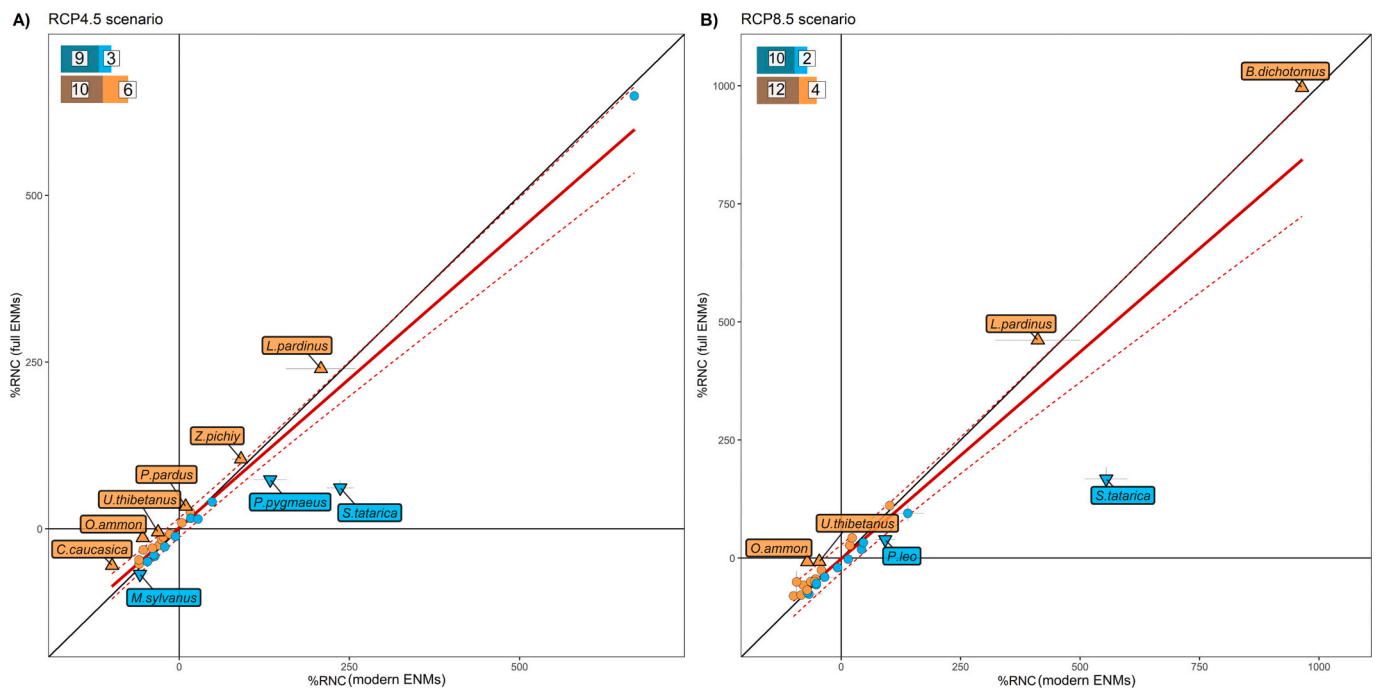


Fig. 1. Range net change values generated for each species by modern (along the x axis) and full (along the y axis) ENMs under RCP4.5 (left) and RCP8.5 (right) scenarios. Orange symbols refer to species reporting a higher RNC in full than in modern ENMs (so-called “higher fossil”), while azure symbols indicate the species showing the opposite pattern (i.e., “higher modern” species). Labeled triangles refer to species that deviate significantly from the mean relationship linking RNC values from modern and full ENMs (the red thick line; dashed lines indicate the 95 % confidence interval). Upper bar plots depict the total number of “higher fossil” (orange) and “higher modern” (azure) species, with the bar portions in brighter colours indicating the number of species significantly deviating from the mean relationship between modern and full ENMs RNC values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Niche overlap analyses

Results for 10 out of 38 species that achieved both modern and full ENMs with $AUC < 0.7$ were excluded from the subsequent analyses. Niche overlap analyses indicated low to moderate overlap values between modern and fossil niches among the analyzed species, reporting Schoener’s D values between 0 for West Caucasian tur (*Capra caucasica*) to 0.51 for reindeer (*Rangifer tarandus*). We found that all the species increased their niche width after fossil data inclusion, with almost a half of them (i.e., 14 out of 28) even doubling it (Fig. S1). The overall average increase is 204 %, with sambar deer (*Rusa unicorn*) showing the lowest increase (i.e., 2 %) and Iberian lynx (*Lynx pardinus*) the highest (i.e., 1586 %).

3.2. Ecological niche models

The 28 species showing AUC values > 0.7 in both modern and full ENMs reported a mean AUC equal to 0.82 ($sd = 0.07$) and a mean CBI equal to 0.67 ($sd = 0.20$) for the former (Table S2), and a mean AUC equal to 0.79 ($sd = 0.06$) and mean CBI of 0.61 ($sd = 0.21$) for the latter (Table S2). These species were included in the further analyses. When validated under temporal block cross-validation, modern ENMs substantially failed in predicting species distribution (mean $AUC = 0.56$, $sd = 0.17$; mean CBI = 0.03, $sd = 0.34$; Table S2), indicating poor temporal transferability.

ENMs predictions under future climate change scenarios revealed similar behaviours between modern and full ENMs in terms of RNC values (Fig. 1; Table S3). At some 6 to 10 times more range projected than currently predicted, the marsh deer (*Blastoceros dichotomus*) consistently is the species with the highest RNC, irrespective of the climate change scenario and the ENM applied. Under RCP4.5,

C. caucasica (−98 % with modern ENM) and Barbary macaque (*Macaca sylvanus*; −67 % with full ENM) are expected to experience the highest contraction in suitable range. Under RCP8.5 scenario, both modern and full ENMs predicted that the worst scenario regards *C. caucasica*, with −100 % and −80 %, respectively (Fig. 1; Table S3).

The slope of the regression between the RNC predicted by modern versus the RNC predicted by full ENMs under RCP4.5 is 0.892 ($p < 0.001$) and the intercept is not significantly different from zero ($p = 0.776$). Under RCP8.5, the slope was 0.876 ($p < 0.001$) and the intercept not significantly different from zero ($p = 0.914$). Of the 56 RNC values (28 species multiplied by 2 ENM types), 55 were congruent (Table S3), meaning that the predicted RNC went in the same direction irrespective of whether fossil data were included or not. These results indicate that the inclusion of fossil data does not change the main insight as per the species fate (i.e., whether the range is expected to decrease or to increase either), rather affecting the magnitude of the predicted range change. In fact, under RCP4.5 scenario, 16 out of 28 species reported higher RNC values from the full than from the modern ENMs. Of these 16 “higher fossil” species, six significantly deviate from the average relationship between RNC scores of modern and full ENMs (i.e., *C. caucasica*, *Ovis ammon*, *Ursus thibetanus*, *Panthera pardus*, *Zaedyus pichiy* and *L. pardinus*; Fig. 1; Table S3). Among the remaining 12 species (i.e., “higher modern”), only three showed significantly higher RNC values from modern than full ENMs (i.e., *M. sylvanus*, *Pongo pygmaeus* and *Saiga tatarica*; Fig. 1). Still under RCP8.5 scenario, 16 species resulted to be “higher fossil” and 12 “higher modern”. Four species show significantly higher RNC values from full than modern ENMs (i.e., *O. ammon*, *U. thibetanus*, *L. pardinus* and *B. dichotomus*), while two report the opposite trend (i.e., *P. leo* and *S. tatarica*; Fig. 1; Table S3).

3.3. Range net change versus niche width gain granted by fossil data

LMM including species outcome as interaction term to explain the

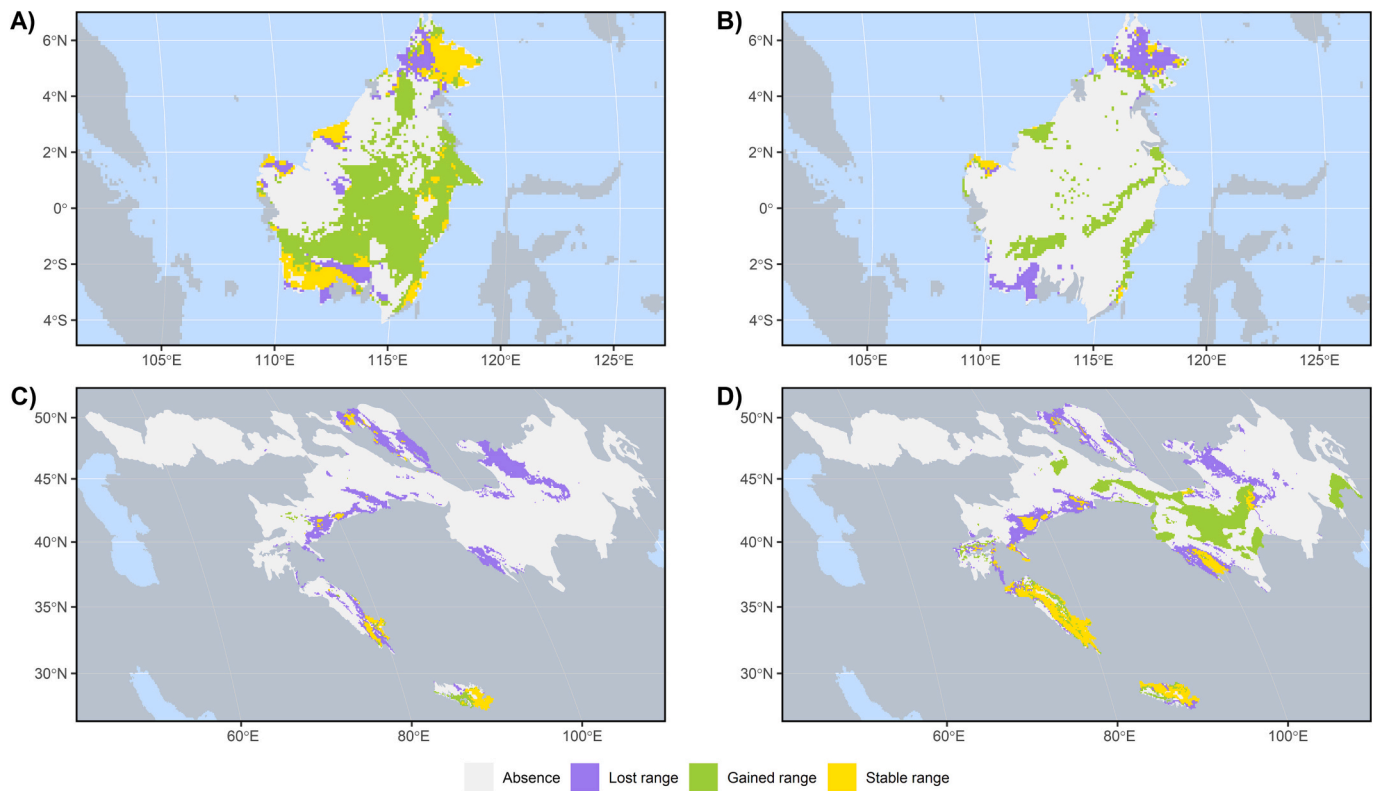


Fig. 2. Spatially explicit predictions of *Pongo pygmaeus* (panels A and B) and *Ovis ammon* (panels C and D) range modifications under 2080 climate change scenarios, as generated by modern (panels A and C) and full (panels B and D) ENMs. Maps refer to the “maximize TSS” binarization threshold. Light grey: stable unsuitable areas; purple: range loss; yellow: stable range; green: range gain. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

relationship between the increase in niche width generated by fossil data and Δ RNC has a stronger statistical support than the other candidate model without this interaction (AIC with species outcome in interaction: 14981.38; AIC without species outcome in interaction: 15748.93). The best LMM achieved a high goodness-of-fit, with conditional R^2 value = 0.46. According to this model, an increase in niche width after fossil data inclusion is significantly correlated with a higher Δ RNC, namely a higher RNC in full than in modern ENMs. However, the sign of this correlation is not univocal among the species, resulting significantly decreasing in “higher modern” species and significantly increasing in “higher fossil” species (Fig. 3; Table S4). This diverging pattern is significantly more pronounced under the RCP8.5 scenario (Fig. 3; Table S4).

3.4. Environmental differences in fossil niche between “higher modern” and “higher fossil” species

Both RF models for RCP4.5 and RCP8.5 scenarios achieved high classification performances, scoring accuracy rates higher than 80 %. According to both RF models, the species with a greater probability to be classified as “higher fossil”, that is to report a higher RNC from full than from modern ENMs, are those where the inclusion of fossil data provide warmer and drier climates to the species niche (Fig. 4). Specifically, the chance for a species to be classified as “higher fossil” strongly increases toward higher temperatures of the warmest month, the wettest season and, to a lesser extent, toward intermediate–high temperature seasonality values (particularly under RCP8.5 scenario). In parallel, the chance for a species to be “higher fossil” strongly decreases toward higher precipitation values during the driest month and season (Fig. 4).

After removing the fossil localities older than the most climatically divergent time intervals, we found a significant reduction in species climate niche width, with respect of the full fossil record ($t = 4.0154$, p

< 0.001). That said, all the insights we gained from using the full record remain unaltered (Supplementary text S1).

4. Discussion

In this study, we provide evidence that pooling modern and fossil data sources to inform climate change vulnerability assessments via ENMs determines a systematic increase in the niche width of the analyzed species and yet modern and full ENMs predicted consistent trends in terms of range increase/decrease. While this result is not unexpected and reported elsewhere (Lima et al., 2017; Sales et al., 2022), we also found the two ENM groups showed diverging predictions about the magnitude of range change, with several species reporting significant discrepancies in the RNC values predicted by the two model classes. This merits further inquiry, especially if designing conservation plans based on ENMs range change predictions is at stake. According to our data, the generalized increase in niche width determined by the inclusion of fossil data does not imply a monotonic improvement in the predicted effects of climate change. In particular, we found that some one half of the species reported lower RNC values under the full ENMs.

There is widespread acknowledgement that fossil data pertaining to living species help understanding their evolution, anatomy and ecology. For what concerns species ecology, fossil information provides valuable data about their area of origin, potential distribution, and reaction to past climatic changes that are hard to gain by relying on the current distribution only (Jones et al., 2019; Maiorano et al., 2013; Mondanaro et al., 2020). For instance, fossil data inform about species ecological plasticity (Di Febbraro et al., 2017), phenotypic adaptation (Mitchell et al., 2018; Smith and Betancourt, 2003), and geographic range shift in response to climate change (Beyer and Manica, 2020; Carotenuto et al., 2016). Since fossil data represent moments of evolution with potentially non-analogue (to the current conditions) climates and are mostly free

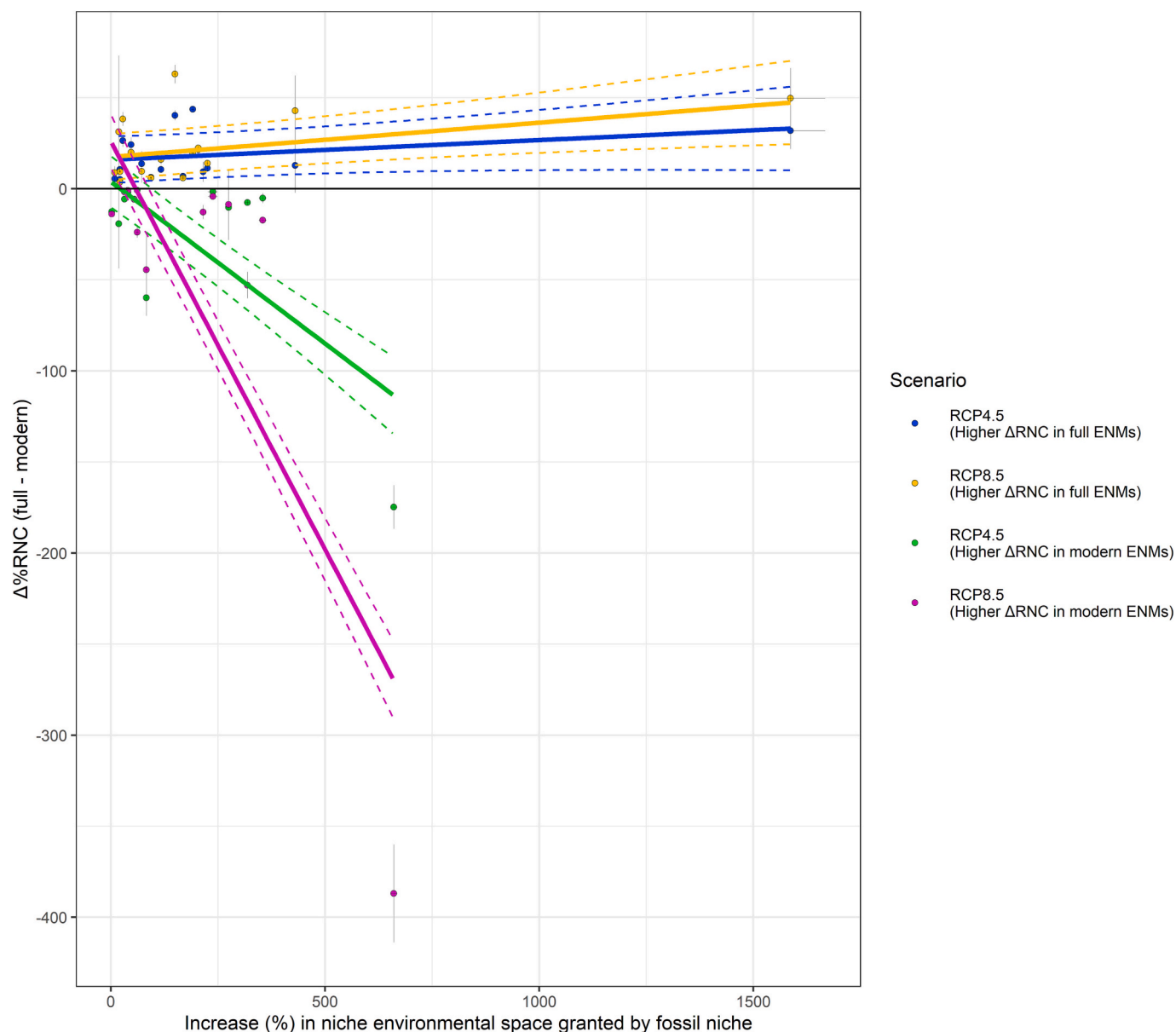


Fig. 3. Linear Mixed Model relationship fitted between the percentage increase in niche environmental space generated by the inclusion of fossil data and the difference in range net change values from full and modern ENMs. Tick lines indicate the statistical relationship between x and y variables, while dashed lines refer to 95 % confidence interval. Each point in the plot represents the mean values of the two above-mentioned variables for each species, while error lines depict standard error values.

from any human impact, our starting hypothesis was that the climatic niche width of species could only increase once their fossil occurrences were considered. In keeping with this hypothesis, we found an average increase in niche width of ca. 180 % when pooling modern and fossil data (Fig. S1), while the degree of niche overlap was generally modest, probably because of the sampling of non-analogue climates. This outcome is in contrast with the substantial evidence for niche stability in corals found by Jones et al. (2019), though this latter research was set to cover a relatively shorter temporal span (i.e., Last Interglacial, ca. 125 kya), compared to ours. Realized niche evolution could be exceedingly fast in modern species and clades (Atwater et al., 2018; Castiglione et al., 2022; Di Marco et al., 2021). Yet, niche conservatism is probably common (Cooper et al., 2011; Liu et al., 2020; Peterson et al., 1999), suggesting that rapid climatic niche evolution might refer to the exploration of different spaces of the fundamental niche, rather than to its evolution (Pearman et al., 2008; Tingley et al., 2014). This is probably the case with our data. Despite the generally low Schoener's *D* values, the results

obtained by modern and full ENMs in terms of RNC trends (range increase vs. decrease) are congruent. This means that modern ENMs are very good predictors of full ENMs results (Fig. 1). That said, the two model classes diverge in terms of RNC magnitude, with a majority of species reporting higher RNC under full than modern ENMs. This may result in substantial differences in terms of where suitable habitats are predicted to occur in the future (Fig. 2), raising concern about the use of modern data only to project species suitable ranges in the future for conservation planning.

Our second hypothesis was that climatic niche expansion by inclusion of fossil data would result in a reduced impact of climate change on the future geographic distribution in most cases. Interestingly, we did not find such a consistent pattern among the mammal species we studied, with almost a half of them showing worsened predicted climate change effects after the inclusion of fossil data. This outcome is in contrast with reports presented elsewhere (Lima et al., 2017; Sales et al., 2022), which tend to suggest that including fossil data determines an

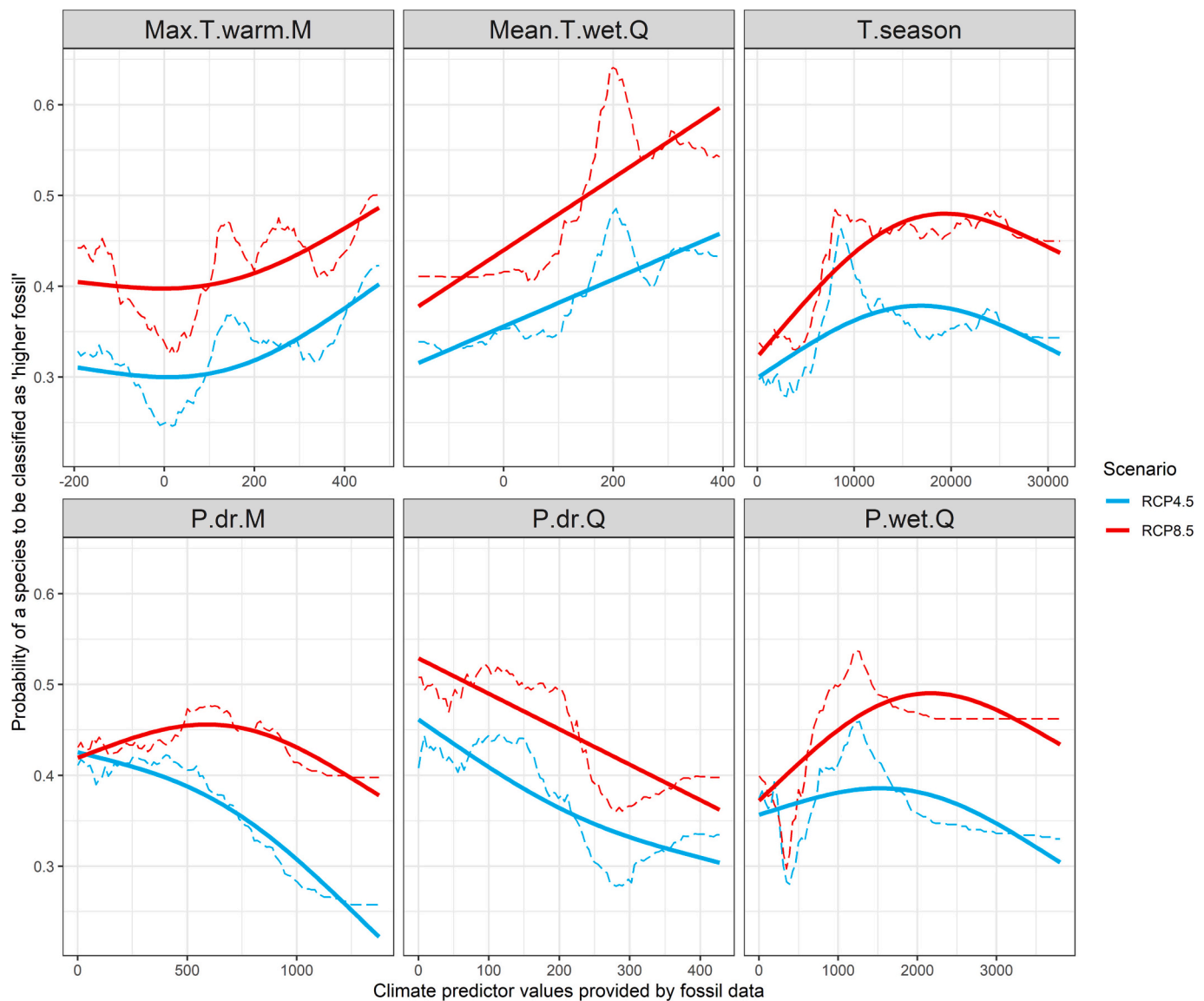


Fig. 4. Dashed lines describe the relationship between climate variable values within the fossil niche and the probability of a species to be classified as “higher fossil”, as generated by random forest model. To improve readability, curves generated by a gam spline were superimposed.

improved reaction to climate change. Sales et al. (2022) reported a prevalent increase in species niche width after the inclusion of fossil data, exactly as we found, though indicating full ENMs to predict higher RNC than modern ENMs in almost 75 % of the analysed species. Moreover, the remaining ca. 25 % species reported a modest discrepancy between modern and full ENMs predictions of range change (i.e., >10–12 % on average). In contrast, we found the species showing higher RNC values according to modern ENMs was as high as some 45 %, reporting discrepancies in range change values >40 % on average. Such divergent outcomes might be related to the different modelling approach deployed here (i.e., a classical ENM framework) as contrasted to Sales et al. (2022), who used reconstructions of past geographic distributions obtained from a mixture of different techniques, instead of explicitly modelling them relying on fossil data. Alternative modelling strategies exist that supposedly allow to account for evolution in estimating niche characteristics. For instance, Smith et al. (2019) suggested that capturing niches at supraspecific taxonomic levels (i.e., “data lumping”) might improve environmental preference estimation. While we cannot exclude that applying such an alternative approach in this context might have led to different outcomes, it is important to mention that data lumping is mostly recommended for related taxa sharing wide

portions of their niches, which is obviously not the case addressed here.

The species we classified as “higher fossil” had experienced past climatic conditions that are expected to be brought about by current climatic change (Fig. 4). The converse seems to be true of the “higher modern” species group. These results, along with the very good agreement between the two ENMs classes, suggest that most species in our data seem to have explored, in the past, different domains of their respective fundamental niches. The idiosyncrasies separating “higher fossil” from “higher living” species groups probably rely in the mean bioclimatic preference of their fundamental niche. The global temperature increases and extensive aridification expected to occur in the next few decades could potentially favour species that were capable to exploit these conditions in the past, regardless of how dry and warm the environment they now occupy are, and the other way around for those which show preference for cooler/wetter conditions during their evolutionary past (Fig. 4). This calls for the importance of considering the fossil record of species occurrences to fully understand their bioclimatic preferences, as several species nowadays possibly are in non-equilibrium with their environments.

In summary, our study suggests that the inclusion of fossil data in calibrating ENMs gives results strongly consistent with the insight

gained at using modern data only, meaning that species will either gain or lose suitable territories irrespective of the ENM approach used. However, the different magnitudes in predicted climate change effects between the two model classes suggest caution in defining conservation actions relying exclusively on ENM calibrated with modern data. For instance, a high vulnerability to climate change predicted for a given species by considering only modern data might suggest that it currently occupies suboptimal climates due to some form of niche truncation. Unravelling the true potential of such a species to face future climate change impacts can only be possible by exploring its realized niche portions expressed in the past, i.e., relying on its fossil record.

The evidence provided in this study highlights the potential for fossil data to open a window into the possible reactions to future threats to species distribution that are not apparent by using modern data only.

Author contributions

AMB, AM & MDF conceived the study. AM, AMB, GG, SC, MM collected the data. AMB, PR, and MDF performed the analyses. All authors contributed significantly to writing the manuscript and preparing submission material.

Funding

This study was supported by the “CitAlien” departmental project funded by the Department of Biosciences and Territory, University of Molise.

CRediT authorship contribution statement

Arianna M. Belfiore: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Alessandro Mondanaro:** Conceptualization, Data curation, Writing – original draft, Writing – review & editing. **Silvia Castiglione:** Data curation, writing - review & editing. **Marina Melchionna:** Data curation. **Giorgia Girardi:** Data curation. **Pasquale Raia:** Conceptualization, Formal analysis, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Mirko Di Febraro:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

Occurrence data, climatic variables and R code are available from the Dryad Digital Repository doi: <https://doi.org/10.5061/dryad.02v6wwq96>.

Acknowledgements

This work was supported by the PNRR DM 351/2022 (M4C1-Inv. 4.1) CUP H53C22000710001 thematic “Innovative forecasting approaches to evaluate climate change effects on biodiversity”; contract number DOT197W57F by the Ministry of University and Research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.biocon.2024.110495>.

References

- Atwater, D.Z., Ervine, C., Barney, J.N., 2018. Climatic niche shifts are common in introduced plants. *Nat. Ecol. Evol.* 2, 34–43. <https://doi.org/10.1038/s41559-017-0396-z>.
- Barbet-Massin, M., Rome, Q., Villemant, C., Courchamp, F., 2018. Can species distribution models really predict the expansion of invasive species? *PLoS One* 13, e0193085. <https://doi.org/10.1371/journal.pone.0193085>.
- Barnett, R., Shapiro, B., Barnes, I., Ho, S.Y.W., Burger, J., Yamaguchi, N., Higham, T.F.G., Wheeler, H.T., Rosendahl, W., Sher, A.V., Sotnikova, M., Kuznetsova, T., Baryshnikov, G.F., Martin, L.D., Harrington, C.R., Burns, J.A., Cooper, A., 2009. Phylogeography of lions (*Panthera leo* ssp.) reveals three distinct taxa and a late Pleistocene reduction in genetic diversity. *Mol. Ecol.* 18, 1668–1677. <https://doi.org/10.1111/j.1365-294X.2009.04134.x>.
- Bartoň, K., 2016. MuMIn: Multi-model inference.
- Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Beyer, R.M., Manica, A., 2020. Historical and projected future range sizes of the world's mammals, birds, and amphibians. *Nat. Commun.* 11, 5633. <https://doi.org/10.1038/s41467-020-19455-9>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Breiner, F.T., Guisan, A., Bergamini, A., Nobis, M.P., 2015. Overcoming limitations of modelling rare species by using ensembles of small models. *Methods Ecol. Evol.* 6, 1210–1218. <https://doi.org/10.1111/2041-210X.12403>.
- Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H., Guisan, A., 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. *Glob. Ecol. Biogeogr.* 21, 481–497. <https://doi.org/10.1111/J.1466-8238.2011.00698.X>.
- Carotenuto, F., Di Febraro, M., Melchionna, M., Castiglione, S., Saggese, F., Serio, C., Mondanaro, A., Passaro, F., Loy, A., Raia, P., 2016. The influence of climate on species distribution over time and space during the late quaternary. *Quat. Sci. Rev.* 149, 188–199. <https://doi.org/10.1016/j.quascirev.2016.07.036>.
- Castiglione, S., Serio, C., Mondanaro, A., Melchionna, M., Carotenuto, F., Di Febraro, M., Profico, A., Tamagnini, D., Raia, P., 2020. Ancestral state estimation with phylogenetic ridge regression. *Evol. Biol.* 47, 220–232. <https://doi.org/10.1007/s11692-020-09505-x>.
- Castiglione, S., Mondanaro, A., Di Febraro, M., Melchionna, M., Serio, C., Girardi, G., Belfiore, A.M., Raia, P., 2022. Testing for changes in rate of evolution and position of the climatic niche of clades. *Mamm. Rev.* 52, 573–583. <https://doi.org/10.1111/mam.12303>.
- Chiarenza, A.A., Waterson, A.M., Schmidt, D.N., Valdes, P.J., Yesson, C., Holroyd, P.A., Collinson, M.E., Farnsworth, A., Nicholson, D.B., Varela, S., Barrett, P.M., 2023. 100 million years of turtle paleoniche dynamics enable the prediction of latitudinal range shifts in a warming world. *Curr. Biol.* 33, 109–121.e3. <https://doi.org/10.1016/j.cub.2022.11.056>.
- Cooper, N., Freckleton, R.P., Jetz, W., 2011. Phylogenetic conservatism of environmental niches in mammals. *Proc. R. Soc. B Biol. Sci.* 278, 2384–2391. <https://doi.org/10.1098/rspb.2010.2207>.
- Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F.T., D’Amen, M., Randin, C., Engler, R., Pottier, J., Pio, D., Dubuis, A., Pellissier, L., Mateo, R.G., Hordijk, W., Salamin, N., Guisan, A., 2017. Ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography* 40, 774–787. <https://doi.org/10.1111/ECOG.02671>.
- Di Febraro, M., Carotenuto, F., Castiglione, S., Russo, D., Loy, A., Maiorano, L., Raia, P., 2017. Does the jack of all trades fare best? Survival and niche width in late Pleistocene megafauna. *J. Biogeogr.* 44, 2828–2838. <https://doi.org/10.1111/jbi.13078>.
- Di Febraro, M., Menchetti, M., Russo, D., Ancillotto, L., Aloise, G., Roscini, F., Preatoni, D.G., Loy, A., Martinoli, A., Bertolino, S., Mori, E., 2019. Integrating climate and land-use change scenarios in modelling the future spread of invasive squirrels in Italy. *Divers. Distrib.* 25, 644–659. <https://doi.org/10.1111/ddi.12890>.
- Di Marco, M., Pacifici, M., Maiorano, L., Rondinini, C., 2021. Drivers of change in the realised climatic niche of terrestrial mammals. *Ecography* 44, 1180–1190. <https://doi.org/10.1111/ecog.05414>.
- Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Syst.* 40, 677.
- Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* 17, 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>.
- Fois, M., Cuenca-Lombrana, A., Fenu, G., Bacchetta, G., 2018. Using species distribution models at local scale to guide the search of poorly known species: review, methodological issues and future directions. *Ecol. Model.* 385, 124–132. <https://doi.org/10.1016/j.ecolmodel.2018.07.018>.
- Franklin, J., Davis, F.W., Ikegami, M., Sypard, A.D., Flint, L.E., Flint, A.L., Hannah, L., 2013. Modeling plant species distributions under future climates: how fine scale do climate projections need to be? *Glob. Chang. Biol.* 19, 473–483. <https://doi.org/10.1111/gcb.12051>.
- Frans, V.F., Augé, A.A., Fyfe, J., Zhang, Y., McNally, N., Edelhoff, H., Balkenhol, N., Engler, J.O., 2022. Integrated SDM database: enhancing the relevance and utility of species distribution models in conservation management. *Methods Ecol. Evol.* 13, 243–261. <https://doi.org/10.1111/2041-210X.13736>.
- Fritz, S.A., Schnitzler, J., Eronen, J.T., Hof, C., Böhning-Gaese, K., Graham, C.H., 2013. Diversity in time and space: wanted dead and alive. *Trends Ecol. Evol.* 28, 509–516. <https://doi.org/10.1016/j.tree.2013.05.004>.

- Slater, G.J., Harmon, L.J., Alfaro, M.E., 2012. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution (NY)* 66, 3931–3944. <https://doi.org/10.1111/j.1558-5646.2012.01723.x>.
- Smith, A.B., Godsoe, W., Rodríguez-Sánchez, F., Wang, H.H., Warren, D., 2019. Niche estimation above and below the species level. *Trends Ecol. Evol.* <https://doi.org/10.1016/j.tree.2018.10.012>.
- Smith, F.A., Betancourt, J.L., 2003. The effect of Holocene temperature fluctuations on the evolution and ecology of *Neotoma* (woodrats) in Idaho and northwestern Utah. *Quatern. Res.* 59, 160–171. [https://doi.org/10.1016/S0033-5894\(03\)00004-8](https://doi.org/10.1016/S0033-5894(03)00004-8).
- Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. *Science* 199 (240), 1285–1293.
- Syfert, M.M., Smith, M.J., Coomes, D.A., 2013. The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PloS One* 8, e55158. <https://doi.org/10.1371/journal.pone.0055158>.
- Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD - a platform for ensemble forecasting of species distribution. *Ecography* 32, 369–373. <https://doi.org/10.1111/j.1600-0587.2008.05742.x>.
- Tingley, R., Vallinoto, M., Sequeira, F., Kearney, M.R., 2014. Realized niche shift during a global biological invasion. *Proc. Natl. Acad. Sci. U. S. A.* 111, 10233–10238. https://doi.org/10.1073/PNAS.1405766111/SUPPL_FILE/PNAS.201405766SI.PDF.
- Valdes, P.J., Armstrong, E., Badger, M.P.S., Bradshaw, C.D., Bragg, F., Crucifix, M., Davies-Barnard, T., Day, J., Farnsworth, A., Gordon, C., Hopcroft, P.O., Kennedy, A. T., Lord, N.S., Lunt, D.J., Marzocchi, A., Parry, L.M., Pope, V., Roberts, W.H.G., Stone, E.J., Tourte, G.J.L., Williams, J.H.T., 2017. The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0. *Geosci. Model Dev.* 10, 3715–3743. <https://doi.org/10.5194/gmd-10-3715-2017>.
- Zuur, A.F., Ieno, E.N., Smith, G.M., 2007. *Analysing Ecological Data*.