

# ECOGRAPHY

## Research article

### Spatially-nested hierarchical species distribution models to overcome niche truncation in national-scale studies

Teresa Goicolea<sup>1</sup>  <sup>1</sup>, Antoine Adde<sup>2,3,4</sup> , Olivier Broennimann<sup>2,3</sup>, Juan Ignacio García-Viñas<sup>5</sup>, Aitor Gastón<sup>5</sup>, María José Aroca-Fernández<sup>5</sup>, Antoine Guisan<sup>2,3</sup>  and Rubén G. Mateo<sup>1</sup> 

<sup>1</sup>Department of Biology, Universidad Autónoma de Madrid, Madrid, Spain

<sup>2</sup>Institute of Earth Surface Dynamics, Faculty of Geosciences and Environment, University of Lausanne, Lausanne, Switzerland

<sup>3</sup>Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

<sup>4</sup>Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

<sup>5</sup>ETSI Montes, Forestal y del Medio Natural, Universidad Politécnica de Madrid, Madrid, Spain

Correspondence: Teresa Goicolea ([teresa.goicolea@uam.es](mailto:teresa.goicolea@uam.es))

#### Ecography

2024: e07328

doi: [10.1111/ecog.07328](https://doi.org/10.1111/ecog.07328)

Subject Editor: Tim Newbold

Editor-in-Chief: Miguel Araújo

Accepted 16 April 2024



[www.ecography.org](http://www.ecography.org)

Spatial truncation in species distribution models (SDMs) might cause niche truncation and model transferability issues, particularly when extrapolating models to non-analog environmental conditions. While broad calibration extents reduce truncation issues, they usually overlook local ecological factors driving species distributions at finer resolution. Spatially-nested hierarchical SDMs (HSDMs) address truncation by merging (a) a global model calibrated with broadly extended, yet typically low-resolution, basic, and imprecise data; and (b) a regional model calibrated with spatially restricted but more precise and reliable data. This study aimed to examine HSDMs' efficacy to overcome spatial truncation in national-scale studies. We compared two hierarchical strategies ('covariate', which uses the global model output as a covariate for the regional model, and 'multiply', which calculates the geometric mean of the global and regional models) and a non-hierarchical strategy. The three strategies were compared in terms of niche truncation, environmental extrapolation, model performance, species' predicted distributions and shifts, and trends in species richness. We examined the consistency of the results over two study areas (Spain and Switzerland), 108 tree species, and four future climate scenarios. Only the non-hierarchical strategy was susceptible to niche truncation, and environmental extrapolation issues. Hierarchical strategies, particularly the 'covariate' one, presented greater model accuracy than non-hierarchical strategies. The non-hierarchical strategy predicted the highest overall values and the lowest decreases over time in species distribution ranges and richness. Differences between strategies were more evident in Switzerland, which was more affected by niche truncation issues. Spain was more negatively affected by climate change and environmental extrapolation. The 'covariate' strategy exhibited higher model performance than the 'multiply' one. However, uncertainties regarding model temporal transferability advocate for adopting and further examining multiple hierarchical approaches. This research underscores the importance of adopting spatially-nested hierarchical SDMs given the compromised reliability of non-hierarchical approaches due to niche truncation and extrapolation issues.

© 2024 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Keywords: climate change, ecological niche truncation, environmental extrapolation, habitat suitability models, nested models, tree species

---

## Introduction

Species distribution models (SDMs) have become invaluable tools for understanding and predicting the potential geographic occurrence of species (Guisan et al. 2017). These models offer valuable information on the ecological niche requirements of organisms by relating species occurrence with environmental variables (Guisan and Zimmermann 2000). This information is crucial for forecasting species response to future environmental changes (Guisan and Thuiller 2005, Booth 2018). Therefore, SDMs have become a key tool for biodiversity conservation planning, anticipating the potential impacts of land use and climate changes or biological invasions and allocating priority areas for conservation and restoration measures (Guisan et al. 2013, Araújo et al. 2019). Ensuring their reliability and accuracy is essential to avoid allocating conservation efforts and resources into suboptimal areas, while neglecting important biodiversity locations (Tulloch et al. 2016, Sofaer et al. 2019).

Setting an adequate spatial extent to calibrate the models is crucial for accurately capturing species–environment relationships and predicting potential species distributions (Thuiller et al. 2004, Chevalier et al. 2021, 2022, Scherrer et al. 2021). However, as conservation decisions and precise sampling campaigns are mostly taken at national, regional, or local levels, SDM studies are in turn rarely conducted at large spatial extents (Dyderski et al. 2017). This could lead to a spatially truncated representation of the range of the species (Scherrer et al. 2021). Spatial truncation may lead to environmental niche truncation when the calibration extent covers just a portion of species' geographical and environmental range (Pearson et al. 2004, Thuiller et al. 2004). Niche truncation implies an incomplete representation of species' ecological niche that underestimates the environmental conditions in which the species could exist. SDMs affected by niche truncation usually result in model transferability issues when projecting them to other periods or areas (Scherrer et al. 2021). Furthermore, geographically-restricted calibration extents are prone to greater dissimilarity between calibrated and projected environmental conditions (i.e., environmental extrapolation), which reduce SDMs transferability and accuracy (Qiao et al. 2019, Charney et al. 2021, Rousseau and Betts 2022) and exacerbate niche truncation issues (Barbet-Massin et al. 2010, Yates et al. 2018, Rousseau and Betts 2022). While some studies recommended limiting SDMs predictions to areas and periods with only analog conditions to the calibration data (Zurell et al. 2012, Rousseau and Betts 2022), this approach can be challenging when forecasting climate change effects, as future conditions may significantly differ from current ones (Williams and Jackson 2007, Owens et al. 2013).

Large calibration extents (i.e. global or continental) capture species' response to a wide range of environmental conditions

(Chevalier et al. 2021, Scherrer et al. 2021, Kindt 2023) and reduce the proportion of non-analog conditions when projected to other areas or temporal periods (Titeux et al. 2017). While using global-level extents would be the optimal solution to tackle niche truncation and extrapolation issues, it is practically unrealistic as precise species observations and high-resolution environmental data for many thematic categories that are relevant for species distribution modelling (e.g. edaphic, hydrologic, land use and cover, etc.) are rarely available over large geographical extents. Using biased or low resolution information available from large-scale dataset can introduce significant noise into models (Keil et al. 2013, Gastón et al. 2017, Moudrý et al. 2023). Furthermore, each environmental covariate influences species distributions at different spatial scales. For example, while climatic covariates are more relevant at global or continental scales, factors such as topography, soil characteristics, or land use and cover drive species distributions at more regional or local extents and finer resolutions (Mackey and Lindenmayer 2001, McGill 2010, Vicente et al. 2014, McGarigal et al. 2016, Mateo et al. 2017). Therefore, relying only on large-scale covariate and species data may disregard some important species distribution drivers and sacrifice the high-resolution and accuracy offered by regional or national databases, restraining detailed ecological understanding (Austin and Van Niel 2011, Mod et al. 2016).

Spatially-nested hierarchical SDMs (hereafter referred to as hierarchical SDMs) have been implemented to overcome the integration of global- and regional-extent datasets by merging SDMs at multiple geographical and resolution levels (Pearson et al. 2004, Mateo et al. 2019a, Guisan et al. unpubl). The benefit of this integration was shown to be two-fold. First, the global model allows covering a broader range of environmental conditions (i.e. ecological niche) and acknowledging wide-scale species–environment relationships by using broadly extended, but unprecise and coarsely-resolved species occurrence and environmental covariate data. Second, the regional model uses spatially-restricted, but more accurate and high resolution data from regional or local datasets to capture the ecological drivers of species distributions at finer grains (Mateo et al. 2019b).

Previous studies have developed several hierarchical strategies for merging global and regional data (Pearson and Dawson 2003, Gastón and García-Viñas 2010, Kéry and Royle 2010, Keil et al. 2013, Petitpierre et al. 2016, Chevalier et al. 2021). Additionally, a recent study has introduced N-SDM, a high-performance end-to-end species distribution modelling pipeline that facilitates adopting and comparing different hierarchical strategies (Adde et al. 2023a). Yet, the differences between available hierarchical strategies and traditional non-hierarchical SDMs have rarely been assessed (Mateo et al. 2019b, Bellamy et al. 2020, Chevalier et al. 2021, 2022), and no study ever focused on comparing the two hierarchical

strategies available in the N-SDM software ver. 1.0 (i.e. ‘covariate’ and ‘multiply’) (Adde et al. 2023a, b). The outcomes from hierarchical and non-hierarchical strategies might vary depending on factors such as the environmental requirements of the species, the size and environmental range of the study area, or the quality of the input data. Therefore, covering multiple study cases is important to comprehend the pros and cons of these approaches, to gain insights into their practical and conceptual differences, and to draw robust conclusions for deciding about the best strategy for predicting species distributions in each specific setting. Addressing these differences into a wide context of cases is pivotal for advancing the application of hierarchical strategies in biodiversity conservation and management, particularly within the context of climate change projections.

In this study, we aimed to strengthen hierarchical modelling understanding in the context of two national-level case studies in Spain and Switzerland. Our objective was to explore the consistency of the results over the different datasets and environmental conditions of the two countries. To offer a comprehensive and transferable framework, we explored the ecological patterns of 108 tree species and the influence of varying greenhouse emission scenarios. More specifically, we aimed to compare SDM results obtained under two hierarchical and a non-hierarchical strategy in terms of niche truncation, environmental extrapolation, model performance, species’ predicted range sizes and shifts, and trends in species richness. By analyzing the consistency of these results over the different study areas, species, and climate conditions, we intended to enrich hierarchical modelling understanding and

optimize the use of hierarchical strategies for biodiversity conservation and management on an international scale.

## Material and methods

### Study areas

Our hierarchical approach integrated models obtained at two spatial levels, with different spatial resolution and extent (Fig. 1): ‘regional’ (250 m resolution and national extent) and ‘global’ (1 km resolution and continental extent). The ‘global’ and ‘regional’ labels were chosen to maintain the terminology of previous studies, including the one introducing the N-SDM software that we used for species distribution modelling (Gallien et al. 2012, Adde et al. 2023a). The regional level was designed to capture local drivers influencing plant species distribution, such as landscape features and microtopography. Two regional study areas were considered: Spain and Switzerland. The global level was the same for both studies and covered completely the European biomes (Olson et al. 2001) and aimed to capture plants’ ecological tolerance to broad-scale drivers, such as macroclimate.

### Species data

We studied all possible native tree species with more than 10 occurrences available at each global and regional levels, for a total of 108 species, including 77 for Spain and 80 for Switzerland, with 49 common to both study areas (see the Supporting information for details on the studied species).

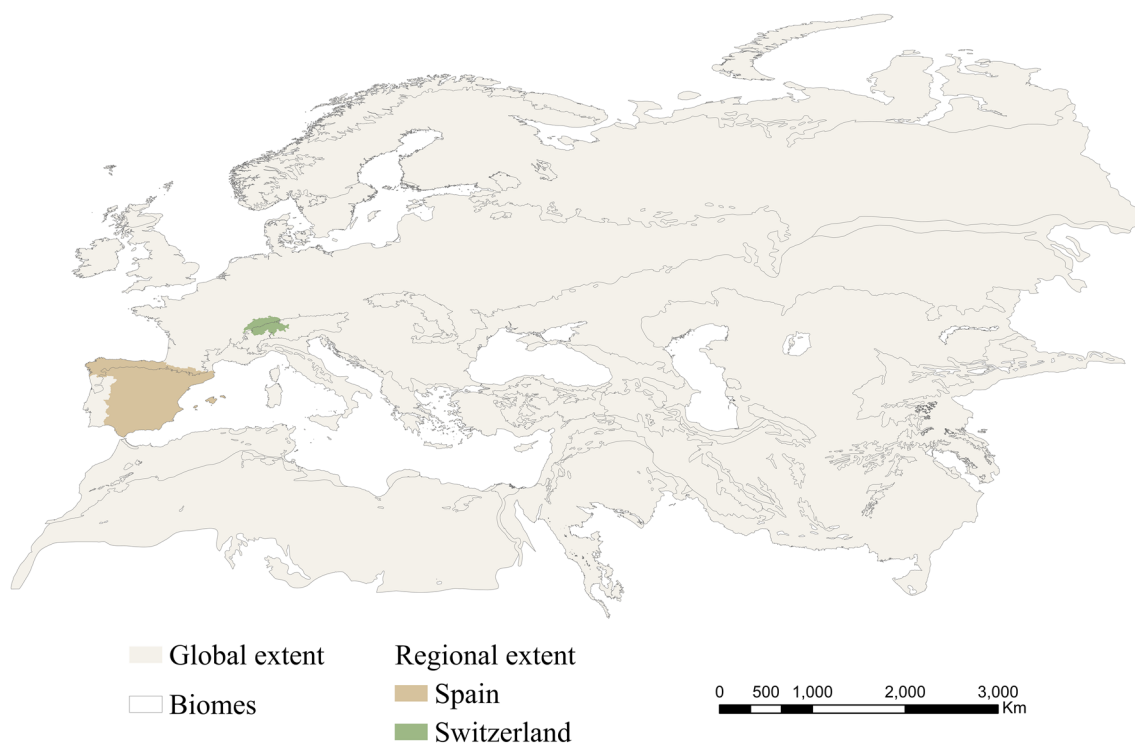


Figure 1. Geographical extent of the global-level and regional-level models.

Species occurrence data for the regional models were obtained from national databases, at a minimum 25 m resolution. Spanish records from 1997 to 2007 were obtained from the third Spanish Forest Inventory (<http://www.mapama.gob.es/es/desarrollo-rural/temas/politica-forestal/inventario-cartografia/inventario-forestal-nacional>), while Swiss records from 1980 to 2021 were obtained from the Species Information Center InfoSpecies ([www.infospecies.ch](http://www.infospecies.ch)). Regional occurrences were disaggregated by removing records closer than 250 m to limit spatial clustering. Global species datasets were obtained from three different sources: GBIF (<https://www.gbif.org/>, DOIs: 10.15468/dl.eedj5f, and 10.15468/dl.4pcvxx), BIEN (<https://bien.nceas.ucsb.edu/bien/>), and EUForest (Mauri et al. 2017). We retained records with at least 1 km resolution and with at least one decimal precision. Expert botanists checked both regional and global occurrences eliminating records that were conclusively identified as erroneous, resulting in an accurate tree presence database. The number of presences removed after expert checking was minimal compared to the remaining number of presences (0.05%).

### Covariate data

We used a set of 27 covariates from four main categories (bioclimatic, edaphic, radiation, and hydrologic) to describe the environmental conditions relevant for plant species distribution (see the Supporting information for details on the covariates). Different resolutions and covariate sets were used for the global and regional models. For the global-level model, we used only the bioclimatic covariates ( $n=19$ ) obtained from the 30-arcsecond resolution climatologies for the Earth's Surface Areas dataset (CHELSA; Karger et al. 2017). For the regional-level models, we used bioclimatic and non-bioclimatic covariates at 250 m resolution. The 19 bioclimatic covariates were downscaled and processed using monthly climatic data from CHELSA. The edaphic covariates (pH, nitrogen content, sand content and soil organic carbon stock at 0–5 cm depth) were obtained from Soilgrids ([www.soilgrids.org](http://www.soilgrids.org)). Hydrologic covariates (distance to rivers, accumulated flow, and topographic index) and the annual solar radiation were calculated from the Digital Elevation Model DAT-193-en (Copernicus Land Cover Service).

To project species' future potential distributions, we retrieved bioclimatic covariates for the 2071–2100 period under four scenarios combining two Global Climatic Models (IPSL\_CM6A\_LR and MRI\_ESM2) and two socio-economic and emissions pathways (SSPs). These SSPs corresponded to optimistic 'low pathway of future greenhouse gas emissions' (SSP126) and pessimistic 'high greenhouse gas emissions' (SSP585).

### Species distribution models

Species distribution models were fitted using the N-SDM software, which is built around a hierarchical modelling framework that allows nesting regional and global level models with two possible approaches ('covariate' and 'multiply'),

and includes several statistical algorithms (Adde et al. 2023a). The high efficiency and parallel processing of N-SDM are useful to simultaneously predict the distributions of multiple species.

We modelled each species under three strategies: 1) a first hierarchical strategy using the 'covariate' nesting approach, which uses the output of a global-level model as an additional covariate for a regional model, 2) a second hierarchical strategy using the 'multiply' nesting approach, which calculates the geometric mean of the independent outputs of global and regional models and 3) a 'regional-only' strategy, using only regional species occurrences and covariate data.

Predictions from the five statistical algorithms available in N-SDM were combined following an ensemble modelling approach: generalized linear model (GLM; McCullagh and Nelder 1989), generalized additive model (GAM; Hastie and Tibshirani 1986, Guisan et al. 2002), maxnet (MAX; Phillips et al. 2006, Elith et al. 2011), random forest (RF; Breiman 2001, Valavi et al. 2021), and light gradient boosted machine (GBM; Ke et al. 2017). N-SDM default values were used for hyperparameter tuning (Adde et al. 2023a). The N-SDM platform includes an automatic covariate selection tool (covsel; Adde et al. 2023b) that selects the best subset of covariates for each species. We parameterized covsel to select a maximum of 10 covariates for the 'regional-only' strategy, and five in each of the two levels of the hierarchical strategies. We generated 10 000' random background absences for each species and level. Spatial projections were computed for the two 250-m resolution grids covering Spain and Switzerland for both current and future conditions.

The N-SDM 'settings.csv' file used to run N-SDM and the ODMAP reporting protocol (Zurell et al. 2020) can be found in the Supporting information.

### Comparing models, study areas and species

To consistently assess the differences in model predictions and related derivatives among three strategies ('covariate', 'multiply' and 'regional-only'), the following metrics were compared over the 108 species, two periods, two emission scenarios, and two study areas:

- Degree of extrapolation. This metric aimed at evaluating the dissimilarity between projected and training bioclimatic conditions. We computed the degree of extrapolation by performing multivariate environmental similarity surfaces (MESS) analyses (Elith et al. 2010) on the bioclimatic covariates with the MESS function of the 'modEvA' R package (Márcia Barbosa et al. 2013). The projected environmental conditions were extracted from the regional extent under future climate conditions. The training environmental conditions were extracted under current climate conditions from 1) the regional extent to measure the extrapolation of the 'regional-only strategy', and 2) the global extent to measure the extrapolation of the hierarchical ones. We then quantified the degree of extrapolation as the percentage of cells in the regional extent future projections with dissimilar environmental



conditions (negative MESS values) to those in the training dataset. This was calculated separately for the four future scenarios and study areas.

- Niche truncation. For each species, this metric aimed at quantifying how the bioclimatic conditions observed at the regional level were representative of those at the global level. We used the ‘unfilling’ metric of the *ecospat.niche.similarity.test* R function (Di Cola et al. 2017) that measures the proportion of the global environmental niche that is not represented by the regional niche (Petitpierre et al. 2012). The global and regional ecological niches were measured according to the two first axes of a principal components analysis of the bioclimatic covariates at species occurrences, at global and regional levels respectively (Broennimann et al. 2012).
- Model performance. To assess the statistical performance of the models, we performed a split-sample cross-validation procedure repeated 100 times and derived five evaluation metrics, including the area under the curve (AUC) (Fan et al. 2006), the continuous Boyce index (CBI) (Hirzel et al. 2006), the maximized Cohen's Kappa coefficient (maxKappa) (Tang et al. 2015), the maximized true skill statistic (maxTSS; (Guisan et al. 2017) and a consensus ‘Score’ metric that averages the AUC’ ( $AUC' = AUC \times 2 - 1$ ), the maxTSS, and the CBI (Adde et al. 2023a). We assessed the statistical significance of the differences in predictive performances between the strategies with a t-test analysis.
- Spatial patterns of species suitability maps. We compared predicted habitat suitability values by 1) contrasting their frequency distribution under current conditions; and 2) pairing values per pixel and calculating the Pearson correlation coefficient between the three strategies and climatic scenarios.
- Species potential distribution. We calculated potential species distribution by binarizing (presence/absence) the continuous suitability maps using the maxTSS threshold (cf. Guisan et al. 2017 for more details on maximization approaches). We compared the spatial patterns and range area of the resulting binary presence/absence maps according to each strategy and current and future climate scenarios. We also calculated the percentage of species whose range area decreased over time according to the different strategies and future scenarios. Additionally, for each species, we classified study area cells according to their evolution from current to future conditions and calculated the percentage of cells of the study area in each class. Cells were classified as lost when they shifted from presence to absence, as gained when shifted from absence to presence, as stable-absence when they remained as an absence, or stable-presence when they remained as a presence.
- Species richness. As a proxy for species richness, we averaged the probabilities of presence maps of all species (Dubuis et al. 2011). We compared the species richness maps visually and based on the Pearson correlation coefficient between the three strategies and all the climatic scenarios. We calculated the species richness for each study

area according to the three strategies, periods and climatic scenarios. To evaluate differences in the magnitude and direction of species richness evolution over time, we calculated the change in species richness from current to future conditions:

$$\Delta SR = \frac{\text{Future species richness} - \text{Current species richness}}{\text{Current species richness}} \times 100.$$

We compared the species richness and change in species richness across strategies by pairing values per pixel and calculating the Pearson correlation coefficient.

## Results

### Degree of extrapolation and niche truncation

The ‘regional-only’ strategy had an average degree of extrapolation of 48% (-SD- 9.87) in Spain and 0.85% (SD 1.37) in Switzerland, on average for the four future climate scenarios (extrapolation results in the Supporting information). In both countries, the degree of extrapolation for the hierarchical approaches was zero.

The average niche ‘unfilling’ metric showed that 13.51% (SD 18.16) of the global niche was not represented in the regional niche in Spain and 20.45% (SD 20.28) in Switzerland (niche ‘unfilling’ results in the Supporting information). The same trend was found for the shared species modelled in both countries, with an average niche ‘unfilling’ metric of 14.69% (SD 6.07) in Spain and 19.65% (SD 17.34) in Switzerland. There were five and nine species with very poorly filled niches at the regional level (‘unfilling’ metric over 50%) in Spain and Switzerland, respectively. Results for individual species can be accessed from the Supporting information.

### Model performance

SDMs were fitted for all 108 species and three strategies. According to all evaluation metrics but CBI - which only had very high values for all strategies - the two hierarchical strategies performed significantly better than the ‘regional-only’ strategy in both countries (Fig. 2). The prediction Score for individual species is provided in the Supporting information. Among the two hierarchical strategies, the ‘covariate’ method had significantly higher model performance values than the ‘multiply’ one, especially in Switzerland. The average Score value of the ‘covariate’, ‘multiply’, and ‘regional-only’ strategies were 0.93, 0.90 and 0.88 (SDs 0.04, 0.04 and 0.06) respectively in Spain, and 0.93, 0.85 and 0.81 (SDs 0.03, 0.05 and 0.07) respectively in Switzerland. We did not find any correlation between the difference in model performance among hierarchical strategies versus the ‘regional-only’ strategy and the niche ‘unfilling’ metric of each species ( $r=0.06$  and  $r=0.16$  for the ‘covariate’ and ‘multiply’ strategies, respectively).

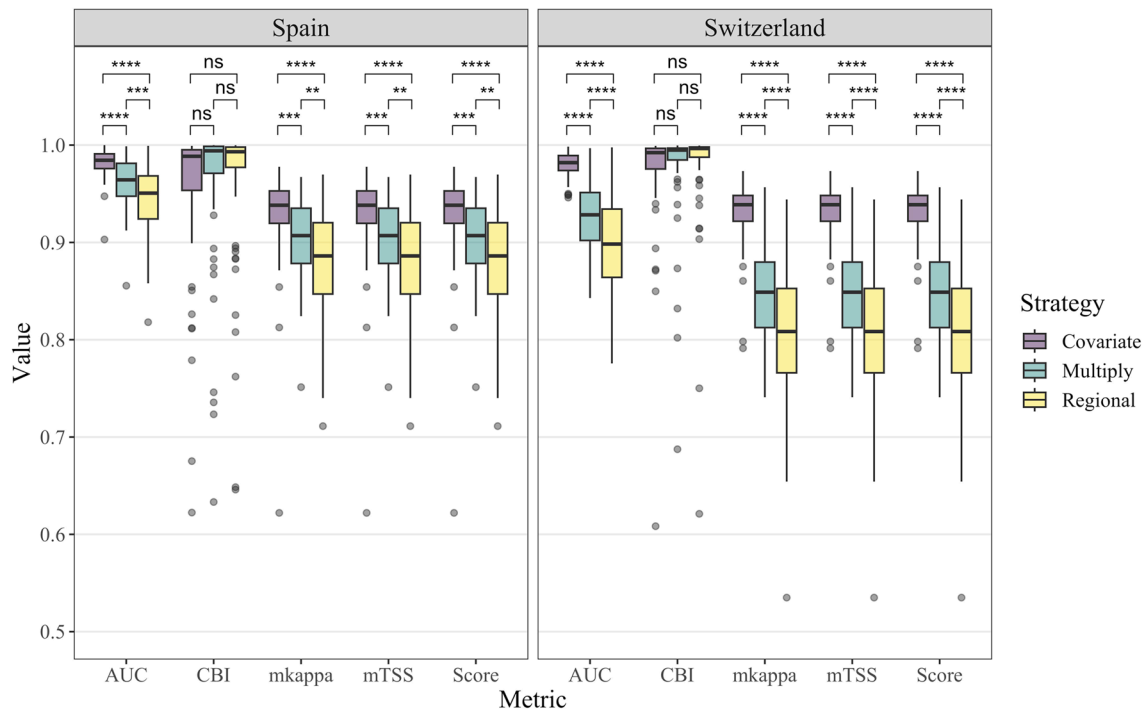


Figure 2. Model performance values obtained for all 108 species in both study areas (Spain and Switzerland), and the three strategies ('covariate', 'multiply', and 'regional-only') from the difference metrics considered: Area under the curve (AUC), continuous Boyce index (CBI), maximized Cohen's Kappa coefficient (mkappa), maximized true skill statistic (mTSS), and 'Score' (average of AUC, mTSS and CBI). Statistical significance of the differences between strategies was assessed using T-tests with \*\*\*\* =  $p < 0.0001$ ; \*\*\* =  $p < 0.001$ ; \*\* =  $p < 0.01$ ; ns = non-significant. In the boxplot, boxes represent the interquartile range of each metric, with the middle horizontal line indicating the median, whiskers show the range of the data excluding outliers, and points depict outliers. Results for individual species can be accessed from the Supporting information.

### Spatial patterns of species suitability maps

Predicted suitability maps were visually similar across strategies, especially in Spain (see an example in the Supporting information). However, the 'covariate' strategy showed more contrasted high and low suitability values, while the 'multiply' and 'regional-only' strategies presented more homogeneous and intermediate suitability values across the study areas (Fig. 3).

The correlation between suitability maps across all species, strategies, and climate scenarios was on average of 0.74 (SD of 0.12) in Spain and of 0.63 (SD of 0.19) in Switzerland (see the Supporting information for correlation and significance values). The average correlation for the species with very high niche truncation ('unfilling' metric over 0.5) was of 0.47 (SD 0.2) in Spain and of 0.58 (SD 0.18) in Switzerland. For a given modelling strategy the suitability maps were also correlated among climate scenarios: in Spain average of 0.74 (SD of 0.13) for the 'covariate' strategy, 0.93 (SD of 0.05) for the 'multiply' strategy, and 0.87 (SD of 0.06) for the 'regional-only' strategy; and in Switzerland average of 0.50 (SD of 0.15) for the 'covariate' strategy, 0.93 (SD of 0.03) for the 'multiply' strategy, and 0.84 (SD of 0.08) for the 'regional-only' strategy.

### Species potential distribution

Spatial patterns in predicted potential distribution (binary maps) were similar across strategies with 90.28% (SD 5.25)

and 82.78% (SD 6.16) of study area cells classified equally as presence/absence across the three strategies in Spain and Switzerland, respectively, under current climate conditions. Species potential distributions were also similarly distributed under future climate conditions across the three strategies. Despite these similarities, the 'covariate' strategy predicted the smallest current range areas in 91.08% of the cases (each case representing a species within a specific study area), while the 'regional-only' the greatest current range areas in 61.1% of the cases. The percentage difference in range area between the 'covariate' and the 'regional-only' strategies was on average -24.54% (SD 20.59) in Spain and -37.30% (SD 18.70) in Switzerland. The 'multiply' strategy generated intermediate predictions that were closer to the 'regional-only' strategy, with a percentage difference in range area with the 'regional-only' strategy of 3.71% (SD 47.64) in Spain and -4.73% (SD 19.79) in Switzerland. In Spain, the differences found between strategies followed a similar trend under current and future climate conditions. In Switzerland, these differences were more pronounced, with larger potential distributions projected in the 'regional-only' strategy compared to the two hierarchical ones. Detailed results on the classification of the pixels as presence or absence across strategies was provided in the Supporting information.

The three strategies predicted range shrinkages for 74.93% (average across all climatic scenarios and strategies, with a SD of 10.17) of species in Spain and for 47.19% (SD of 19.91)

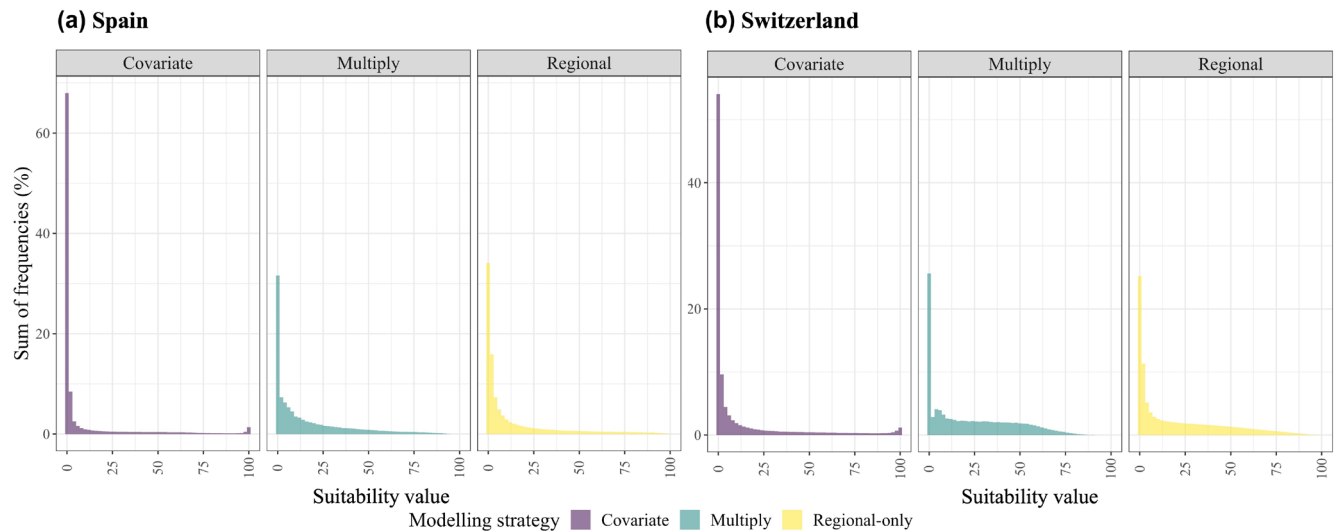


Figure 3. Average histogram frequencies (%) of suitability values for all species under current climate conditions. Histograms were computed separately for Spain and Switzerland, and for the three strategies ('covariate', 'multiply' and 'regional-only').

of species in Switzerland (details in the Supporting information). However, predictions varied among the strategies; both hierarchical strategies predicted a higher number of species with contracted ranges compared to the 'regional-only' approach, particularly in Switzerland. The same trends were found for the 49 shared species modelled in both countries. On average, 75.17% (SD of 8.87) of these shared species experienced range shrinkages in Spain, while 43.3 7% (SD of 22.12) in Switzerland.

Among the three strategies, the 'regional-only' one predicted the smallest ranges shrinkages (the lowest percentage of cells classified as lost), and the largest ranges increases (the biggest percentage of cells classified as gained) (Fig. 4). The 'regional-only' strategy also predicted the biggest percentage of stable-presence and the lowest percentage of stable-absence cells. Among the two hierarchical strategies, the 'covariate' one predicted a greater percentage of gained cells, while the 'multiply' one a greater percentage of stable-presence cells.

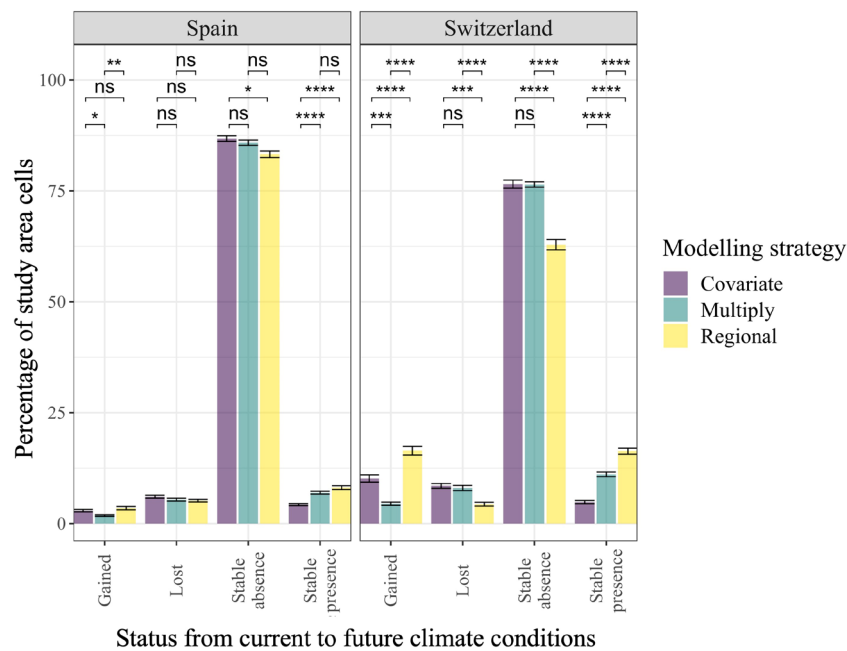


Figure 4. Average percentage of cells classified as lost, gained, stable-presence, or stable-absence from current to future climate conditions (climatic scenario IPSL\_CM6A\_LR\_2070\_SSP585) according to the three strategies ('covariate', 'multiply' and 'regional-only'), and for both study areas (Spain and Switzerland). Error bars indicate the standard errors, representing the variability among species. See the same figure for other climatic scenarios in Supporting information. T-tests statistical significance of the differences between strategies were provided with \*\*\*\* =  $p < 0.0001$ ; \*\*\* =  $p < 0.001$ ; \*\* =  $p < 0.01$ ; \* =  $p < 0.05$ ; ns: non-significant.

These results were consistent across all climate scenarios and study areas.

## Species richness

The mean species richness was significantly different across strategies ( $p < 0.0001$ ), being smaller according to the 'covariate' strategy, followed by the 'multiply' and the 'regional-only' in both countries and under current and all future scenarios (Fig. 5–6). The mean species richness was greater in Switzerland than in Spain. Even though the overall patterns of species richness visually differed according to the

strategy (Fig. 5–6), they remained correlated, with average correlation values between strategies and climate scenarios of 0.92 (SD 0.05,  $p < 0.0001$ ) and 0.86 (SD 0.11,  $p < 0.0001$ ) in Spain and Switzerland, respectively (correlation values in the Supporting information). The areas with the greatest species richness were located in similar regions over time and strategies (mountainous areas and the North range in Spain, and low flat, valley areas and low mountains in Switzerland).

The mean change in species richness over time was significantly different ( $p < 0.0001$ ) between strategies in both countries (Fig. 5–6, Supporting information). In Spain, an overall decrease in species richness was predicted over time consistently

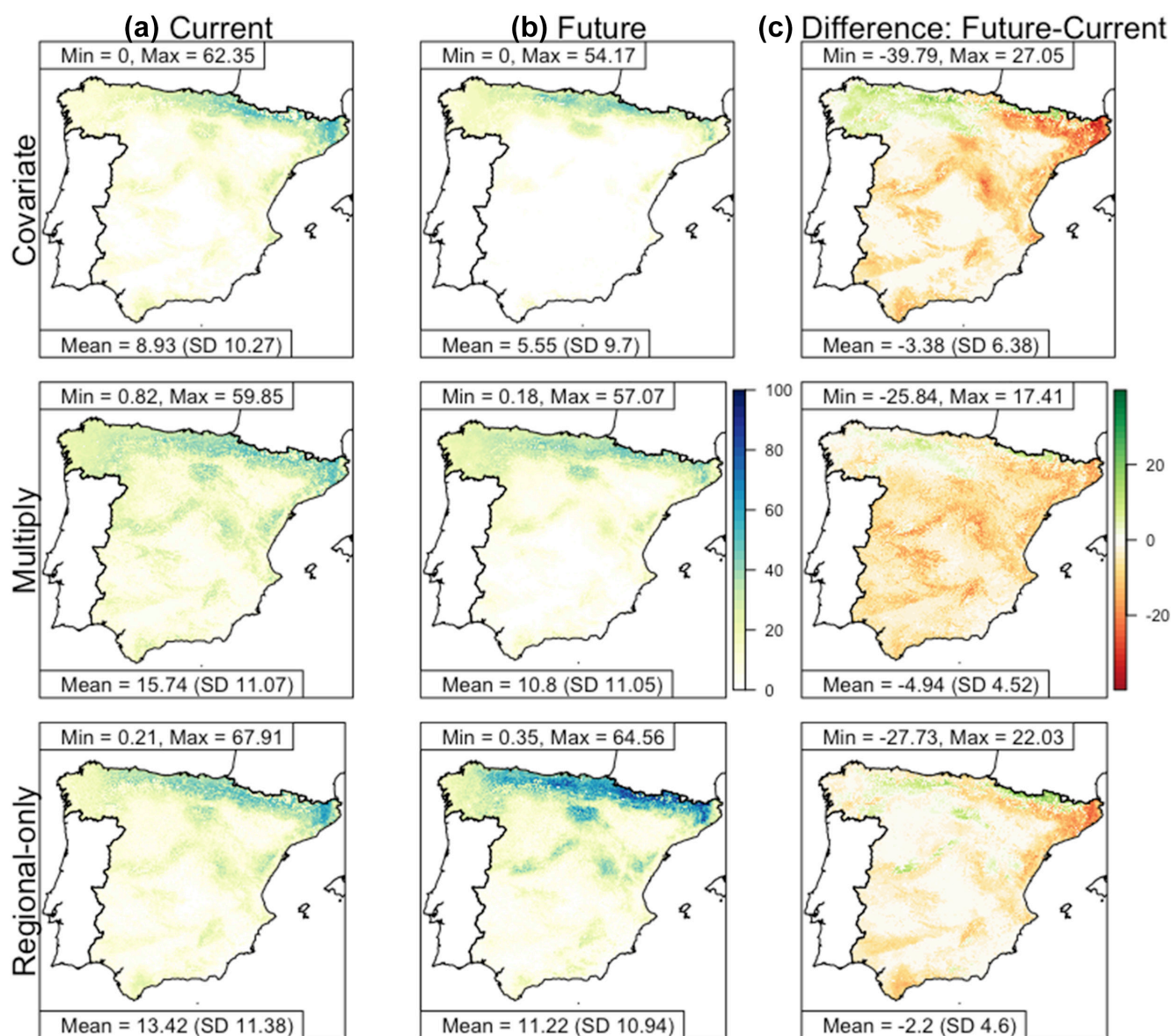


Figure 5. Predicted species richness in Spain under (a) current and (b) future climatic conditions for the IPSL\_CM6A\_LR\_2070\_SSP585 climatic scenario, and (c) difference in estimated species richness between current and future periods according to the three strategies ('covariate', 'multiply' and 'regional-only'). Species richness was estimated by averaging the probability of presence of all species. See the difference in species richness for the other climatic scenarios in the Supporting information. Labels show the minimum, maximum, mean and standard deviation of the values.



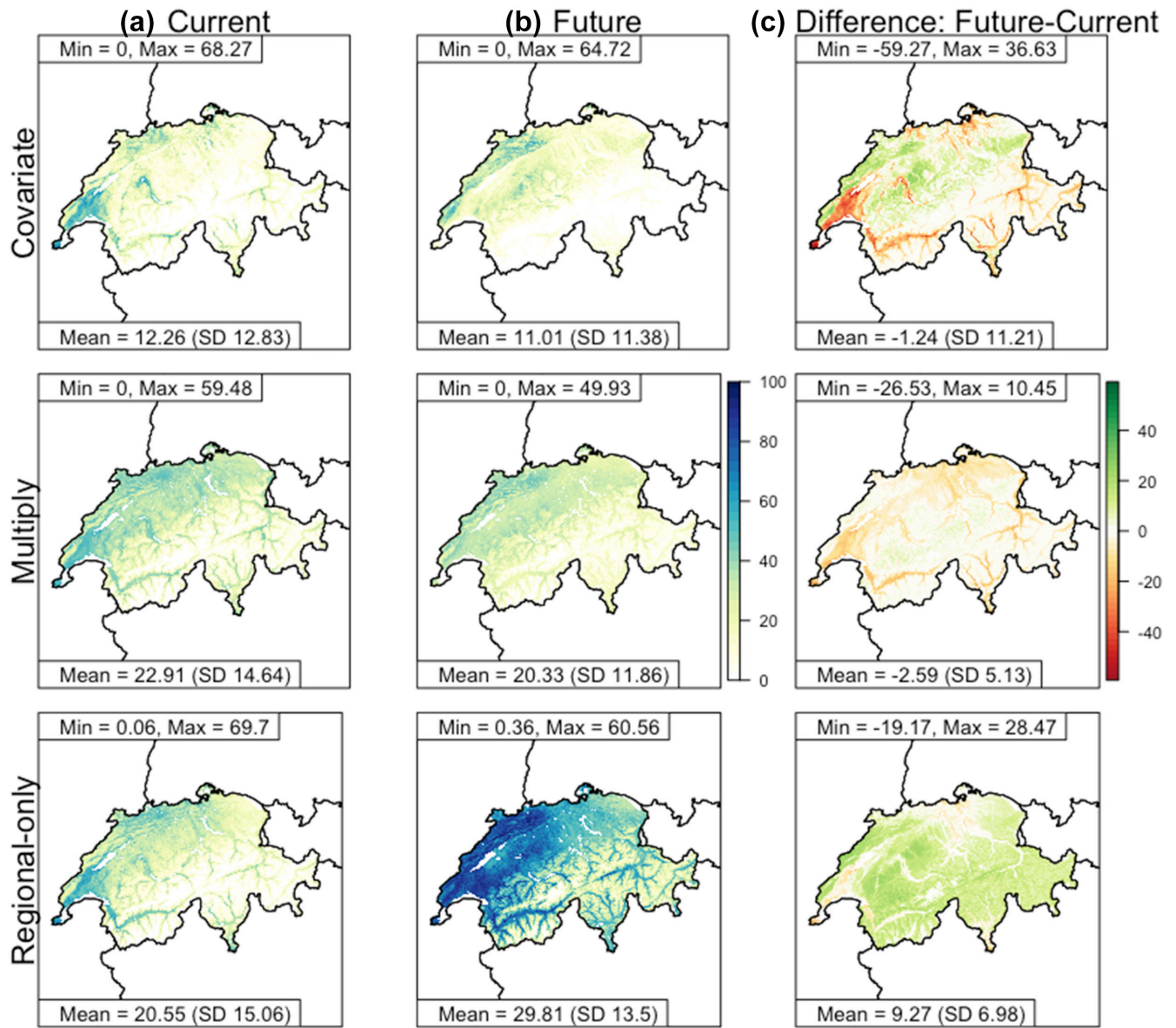


Figure 6. Predicted species richness in Switzerland under (a) current and (b) future climatic conditions for the climatic scenario IPSL\_CM6A\_LR\_2070\_SSP585, and (c) difference in species richness between current and future periods according to the three strategies ('covariate', 'multiply' and 'regional-only'). Species richness was estimated by averaging the probability of presence of all species. See the difference in species richness for the other climatic scenarios in the Supporting information. Labels show the minimum, maximum, mean and standard deviation of the values.

for the three strategies. This decrease was  $-2.18\%$  (SD 0.85),  $-3.24\%$  (SD 1.25) and  $-1.05\%$  (SD 1.07) according to the 'covariate', 'multiply' and 'regional-only' strategies, respectively, on average across all the climate scenarios. In Spain, the proportion of cells with predicted species richness losses was on average 77.99, 85.56 and 64.6 (SD 4.94, 5.36 and 16.62) across all the climate scenarios for the 'covariate', 'multiply' and 'regional-only' strategies respectively. In Switzerland, higher species richness increases were predicted, especially according to the 'regional-only' strategy. On average across all the climate scenarios, the species richness change in Switzerland was  $-0.1\%$  (SD 1.96),  $-1.1\%$  (SD 1.25) and  $6.8\%$  (SD 1.76) according to the 'covariate', 'multiply' and 'regional-only'

strategies. The proportion of study area cells with predicted species richness loss was on average 44.59, 50.64 and 8.07 (SD 8.62, 11.70 and 2.95) across all the climate scenarios for the 'covariate', 'multiply' and 'regional-only' strategies respectively. The spatial patterns of gain (positive values) and loss (negative values) were similar and correlated across the different climate scenarios and strategies. In Spain, there were greater losses in mountainous areas, and the average Pearson correlation across scenarios was of 0.72 (SD 0.16,  $p < 0.0001$ ). In Switzerland, there were greater losses around low elevation areas while species richness stability or increments were projected in mountainous areas, and the average correlation across scenarios was of 0.62 (SD 0.19,  $p < 0.0001$ ).

## Discussion

This study delved into the potential benefits of using spatially-nested hierarchical SDMs by comparing model performances, spatial patterns, species ranges and richness, environmental extrapolation and niche truncation. We highlighted the importance of hierarchical strategies, especially when projecting to environmental conditions that are non-analog to those in the region where the model was calibrated. While the non-hierarchical strategy tested in our study potentially compromised prediction reliability due to niche truncation and extrapolation issues, the two hierarchical strategies considered helped solving these issues, but presented nuanced results. The variety of cases we considered in our study (species, study areas and climate scenarios) strengthens the consistency of our results.

### Predictive performance

Evaluation metrics were higher for hierarchical strategies – particularly for the ‘covariate’ one – over the ‘regional-only’ strategy (Table 1). Previous studies already showed a superior predictive performance of hierarchical models compared to single-level strategies (Bellamy et al. 2020, Chevalier et al. 2021). This superiority is noteworthy as final predictions were assessed using the regional-level species occurrence data (unaffected by niche truncation or extrapolation issues), which could have favored the ‘regional-only’ strategy. The higher performances of hierarchical approaches could be explained by their ability to consider covariate data and species-covariate relationships at relevant spatial scales and resolutions (Mackey and Lindenmayer 2001, Pearson and Dawson 2003, Austin and Van Niel 2011).

The superior model performance of the ‘covariate’ strategy, in comparison to the ‘multiply’ strategy, might also suggest a more effective integration of the global and regional levels by

the former (Bellamy et al. 2020). Indeed, by using the global model output as a covariate, the covariate approach could help obtaining a more accurate estimate of the contribution of the global climatic niche in the final hierarchical model. On the contrary, the ‘multiply’ strategy, by computing the geometric mean of the global and regional model outputs, imposes an equal contribution of the global climatic niche compared to the other covariates, regardless of the species and its relationship with these covariates.

One must be cautious to rely on predictive performance metrics to compare the effectiveness of modelling strategies, particularly when forecasting species occurrence to future periods. The performance metrics used in this study were computed for the current period, and therefore, they do not inherently evaluate the effectiveness of the models to predict at other periods. Ideally, for a comprehensive assessment of the temporal transferability, the training and testing datasets should be temporally independent and collected at time periods with distinct climate conditions (Araújo et al. 2019), although in the latter case, changes in the realized niche might blur the transferability assessment (Maiorano et al. 2013). Nevertheless, the evaluation of future projections is limited by the impossibility to have precise estimates of the future species distributions. Alternatively, SDMs temporal transferability could be evaluated using species occurrences from past periods (Araújo et al. 2005, Pearman et al. 2008, Moreno-Amat et al. 2015), or virtual species (Hirzel et al. 2006, Zurell et al. 2010, Leroy et al. 2016, Chevalier et al. 2021).

### Truncation and extrapolation affect predictions of the ‘regional-only’ strategy

Interestingly, while ‘regional-only’ strategies could be naively expected to underestimate the suitable climatic niche of

Table 1. Summary of the results obtained for the two hierarchical ‘covariate’ and ‘multiply’, and the ‘regional-only’ strategies in Spain and Switzerland for current and future periods. Results for all criteria were averaged across all study tree species, and results for criteria 3, 5, 7 and 8 were also averaged across future scenarios.

Criteria	Study area	Covariate	Multiply	Regional-only
1. Model performance (‘score’)	Spain	0.93	0.90	0.88
	Switzerland	0.93	0.85	0.81
2. Distribution of suitability values	Both countries	Contrasted high or low values	<i>Intermediate and low values</i>	
3. Correlation in suitability values across strategies	Spain	0.74		
	Switzerland	0.63		
4. Mean range area (current period, km <sup>2</sup> )	Spain	50 467	60 393	65 202
	Switzerland	5345	7650	8295
5. % of species with range loss	Spain	76.62	79.87	65.58
	Switzerland	56.56	62.81	22.19
6. Species richness (current period)	Spain	8.93	15.74	13.42
	Switzerland	12.26	22.91	20.55
7. Correlation in species richness across strategies	Spain	0.92		
	Switzerland	0.86		
8. % cells with loss of species richness (future)	Spain	77.99	85.56	64.60
	Switzerland	44.59	50.64	8.07
9. Degree of extrapolation	Spain	None		48.00
	Switzerland			0.85
10. Niche truncation (%)	Spain	None		13.51
	Switzerland			20.45

species and lead to smaller predicted future distributions than models fitted at wider extents (Barbet-Massin et al. 2010, Mateo et al. 2019b, Scherrer et al. 2021), our findings showed the contrary, consistently with other studies (Thuiller et al. 2004, Keil et al. 2013, Bellamy et al. 2020). We found that the ‘regional-only’ strategy tended to over-predict species’ current potential distribution compared to hierarchical ones (Table 1). The over- or under-estimation of predicted distributions is primarily linked to the way the niche is truncated, combined with the extent of environmental extrapolation (i.e. predictions in environments non-analog to the training area), which requires assessing how the shapes of the species’ responses to the environmental predictors are modified by considering a too small extent (Thuiller et al. 2004, Hannemann et al. 2016, Chevalier et al. 2021).

First, the shape of the response curves is highly related to the extent of the ecological range included in the training data (Thuiller et al. 2004). Niche truncation might limit the accurate estimate of the tail of the response curve if there is not enough data to detect a decrease in species presence. For instance, truncated models might fail to detect a unimodal response and substitute it with an asymptotic one (Fig. 7B), or result in other types of transformation (Fig. 1 in Chevalier et al. 2021). The substantial differences between strategies in Switzerland (Table 1), the study area with greatest niche truncation, suggested the essential role of hierarchical approaches in mitigating niche truncation issues (Chevalier et al. 2021). This was underscored by the significant decrease in the correlation between hierarchical and ‘regional-only’ strategies for species exhibiting high niche

truncation. While hierarchical strategies capture a wider range of environmental conditions and mitigate niche truncation, they can still be susceptible to this issue in the case of unrealized niche influenced by factors such as human disturbances, or dispersal limitations (Normand et al. 2011).

Secondly, the shape of the tail of the response curve at the limit of the training data deeply influences the predictions when extrapolating to novel climate conditions (Thuiller et al. 2004, Webber et al. 2011, Chevalier et al. 2021). In this study, in the absence of extrapolation clamping techniques, estimated species response to novel conditions could be amplified when the response curve grew or maintained a consistently elevated probability of presence as environmental conditions approached the limit of the training data (Fig. 7; Owens et al. 2013). Similarly, the species response could be diminished when the response curve decreased or maintained a low value when approaching the limits of the training data. Consequently, the ‘regional-only’ modelling strategy can both overestimate or underestimate species relationship with environmental covariates and therefore the overall species predicted distribution.

### Selecting the best strategy

Niche truncation, environmental extrapolation, and predictive performance results highlighted the benefits of using hierarchical strategies over the ‘regional-only’ one for predictions and management decisions. In this study, we showed that the ‘regional-only’ strategy was overestimating the current and future status of species’ distribution and species

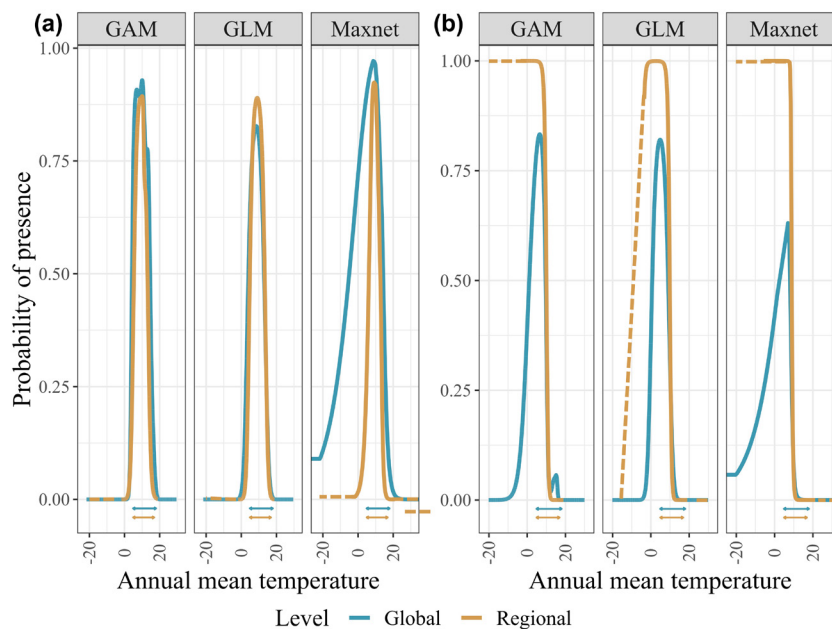


Figure 7. Response curves modelled at global (within hierarchical strategies) and regional levels for (a) *Quercus pyrenaica* (species with low niche truncation at the regional level) and (b) *Juniperus sabina* (species with high niche truncation) in Spain for the covariate ‘annual mean temperature’ for generalized additive model (GAM), generalized linear model (GLM), and Maxnet modelling algorithms. The arrows indicate the annual mean temperature range at the global and regional species presence occurrences. Dashed lines in the regional response curve indicate the extrapolation domain with annual mean temperature values outside the regional training range.

richness. Such overestimations may inadvertently support lower-priority conservation measures compared to those derived from more accurate modeling approaches.

Among the two hierarchical strategies evaluated, the 'covariate' strategy exhibited higher model performances. However, in the 'covariate' strategy, if the global and regional bioclimatic conditions extracted at species occurrence locations are very similar (i.e. low niche truncation), the global-level model output used as a covariate in the regional-level model might over effectively predict species' occurrence probability. It would also have a disproportionately high influence in the model compared to other regional-level covariates. This higher influence could explain why the distribution of the suitability values resulting from the 'covariate' strategy tended to concentrate at the lowest and highest limits (Fig. 3). In contrast, the 'multiply' strategy may reveal a more balanced – yet arbitrary – assignment of covariate contributions, with more intermediate suitability values. In this strategy, the climate suitability is likely more balanced with the edaphic, hydrologic, and topoclimatic covariates. We encourage examining both strategies for fully informed conservation decisions. Additionally, it would be interesting to extend this comparison to other new or already proposed hierarchical strategies (Pearson et al. 2004, Lomba et al. 2010, Mateo et al. 2019a, Chevalier et al. 2022) to better understand their conceptual and practical differences.

### Exceptions to hierarchical strategies

Hierarchical strategies might seem less relevant for species with geographically-restricted distributions or with their entire ecological niche fulfilled at limited extensions (Scherrer et al. 2021). For instance, extrapolating beyond training conditions might be less concerning when the entire species environmental niche is captured (Fig. 7). However, climatic factors influence species distribution at larger scales and coarser resolutions than other habitat covariates (Mackey and Lindenmayer 2001, McGill 2010). Therefore, incorporating these climatic predictors at a coarser resolution than other covariates (e.g. soil, land use and cover, etc.) might help to obtain better predictions, advocating for hierarchical strategies even for species unaffected by niche truncation.

Finally, it is important to be cautious when applying hierarchical strategies to species with different ecotypes, provenances, or subspecies that are not present in the target regional-level area (Scherrer et al. 2021). Incorporating all ecotypes and populations in the global level may result in an overestimation of the species' climatic niche and potential distribution, as ecotypes or subspecies in the regional study area may have restricted niches and local adaptations (e.g. due to dispersal limitations).

### National and ecological insights into key findings

Despite similar trends in results between both countries, the differences among strategies were more pronounced in Switzerland, particularly for future climate conditions. The

more limited climatic extrapolation identified in Switzerland under future climate conditions could be attributed to the significant altitudinal range of this mountainous country. Despite its small size, Switzerland offers a great variability of climatic conditions (Rousseau and Betts 2022) that theoretically allows for a more comprehensive representation of climatic conditions. This characteristic may, in theory, result also in fewer issues related to 'niche truncation' during model calibration in the 'regional-only' strategy, especially for mountain species. Furthermore, the broad range of climate conditions in Switzerland might facilitate altitudinal species distribution shifts, typically moving upwards to access suitable future climate conditions (Hickling et al. 2006, Carroll et al. 2018). This might explain the increase in species richness predicted in mountainous areas (Fig. 6) and the low proportion of species with range shrinkages, as also shown for bird species in Switzerland by Chevalier et al. (2022). On the other hand, we predicted a decrease in species richness under future climate conditions in most Spanish mountainous areas, except for the Cantabrian range (Fig. 5). This finding might indicate that mountainous areas currently represent the climatic limit of suitable conditions for many species, and future climate conditions beyond that limit would not allow those species to persist any longer. In fact, the hotter and drier climate conditions in future forecasts will probably have much more negative effects in Spain than in Switzerland (Loarie et al. 2009). This susceptibility of Spain to climate change was evinced by the great decrease in species range and richness predicted in this study. Beyond these country-specific trends, varying outcomes among species could also highlight the influence of specific individual traits on our findings, being a compelling focus for future research.

### Conclusions

In conclusion, this research emphasizes the importance of adopting spatially-nested hierarchical strategies over 'regional-only' approaches when predicting species distributions, especially if the objective is to obtain temporal or spatial projections. 'Regional-only' strategies may compromise the accuracy of the results due to potential niche truncation and environmental extrapolation issues. Among the two hierarchical strategies we tested, the 'covariate' strategy exhibited higher model performance. However, given uncertainties in future model transferability, we encourage considering multiple strategies for fully-informed conservation decisions.

**Acknowledgements**– We acknowledge the Swiss Species Information Center InfoSpecies ([www.infospecies.ch](http://www.infospecies.ch)) for supplying Swiss-level species occurrence data. We also gratefully acknowledge the support provided by the Scientific Computing and Research Unit ([www.unil.ch/ci/dcsr](http://www.unil.ch/ci/dcsr)). and the computing resources of the HPC cluster of Lausanne University. We are grateful to Jaime Boyano for downloading the species and covariate data.

**Funding**– This study was supported through the Connect2restore project (TED2021-129589B-I00) funded



by Ministerio de Ciencia e Innovación (Agencia Estatal de Investigación) and “Unión Europea NextGenerationEU/PRTR” and the NextDive Project (PID2021-124187NB-I00) funded by Ministerio de Ciencia e Innovación (Agencia Estatal de Investigación) and “FEDER Una manera de hacer Europa” to RGM. AGu benefitted support from the Swiss Federal Office of the Environment (Valpar.CH project).

## Author contributions

Antoine Guisan and Rubén G. Mateo share last authorship. **Teresa Goicolea:** Conceptualization (equal); Data curation (equal); Methodology (equal); Validation (equal); Visualization (lead); Writing – original draft (lead); Writing – review and editing (lead). **Antoine Adde:** Conceptualization (equal); Formal analysis (lead); Methodology (equal); Software (equal); Supervision (equal); Writing – review and editing (lead). **Olivier Broennimann:** Conceptualization (equal); Data curation (equal); Methodology (equal); Writing – review and editing (equal). **Juan Ignacio García-Viñas:** Conceptualization (equal); Data curation (equal); Writing – review and editing (equal). **Aitor Gastón:** Conceptualization (equal); Data curation (equal); Writing – review and editing (equal). **María José Aroca-Fernández:** Conceptualization (supporting); Data curation (supporting); Resources (supporting); Writing – review and editing (equal). **Antoine Guisan:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); Resources (lead); Supervision (lead); Writing – review and editing (equal). **Rubén G. Mateo:** Conceptualization (equal); Formal analysis (lead); Funding acquisition (lead); Project administration (equal); Resources (equal); Supervision (lead); Writing – review and editing (equal).

## Transparent peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ecog.07328>.

## Data availability statement

Climate and edaphic data can be accessed from their source repositories, <https://chelsa-climate.org/> and <https://soilgrids.org/>, respectively. The Digital Elevation Model can be downloaded from Copernicus Land Cover Service. Global-level species data can be accessed from GBIF (<https://www.gbif.org/>, DOIs: [10.15468/dl.eedj5f](https://doi.org/10.15468/dl.eedj5f), and [10.15468/dl.4pcvxx](https://doi.org/10.15468/dl.4pcvxx)), BIEN (<https://bien.nceas.ucsb.edu/bien/>), and EUForest (Mauri et al. 2017). Regional-level species data are available from the third Spanish Forest Inventory (<http://www.mapama.gob.es/es/desarrollo-rural/temas/politica-forestal/inventario-cartografia/inventario-forestal-nacional/>) for Spain and from InfoSpecies ([www.infospecies.ch](http://www.infospecies.ch)) for Switzerland; restrictions apply to the availability of these data, which were used under license for this study. The species distribution modelling software N-SDM is available from the GitHub Repository, <https://github.com/N-SDM/N-SDM>,

where complementary instructions for installation and example data are provided. The N-SDM ‘settings.csv’ file used to run N-SDM can be found in Supporting information.

## Supporting information

The Supporting information associated with this article is available with the online version.

## References

- Adde, A., Rey, P., Brun, P., Külling, N., Fopp, F., Altermatt, F., Broennimann, O., Lehmann, A., Petitpierre, B., Zimmermann, N. E., Pellissier, L. and Guisan, A. 2023a. N-SDM: a high-performance computing pipeline for nested species distribution modelling. – *Ecography* 2023: e06540.
- Adde, A., Rey, P. L., Fopp, F., Petitpierre, B., Schweiger, A. K., Broennimann, O., Lehmann, A., Zimmermann, N. E., Altermatt, F., Pellissier, L. and Guisan, A. 2023b. Too many candidates: embedded covariate selection procedure for species distribution modelling with the covsel R package. – *Ecol. Inform.* 75: 102080.
- Araújo, M. B., Whittaker, R. J., Ladle, R. J. and Erhard, M. 2005. Reducing uncertainty in projections of extinction risk from climate change. – *Global Ecol. Biogeogr.* 14: 529–538.
- Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., Garcia, R. A., Guisan, A., Maiorano, L., Naimi, B., O’Hara, R. B., Zimmermann, N. E. and Rahbek, C. 2019. Standards for distribution models in biodiversity assessments. – *Sci. Adv.* 5: eaat4858.
- Austin, M. P. and Van Niel, K. P. 2011. Improving species distribution models for climate change studies: variable selection and scale. – *J. Biogeogr.* 38: 1–8.
- Barbet-Massin, M., Thuiller, W. and Jiguet, F. 2010. How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? – *Ecography* 33: 878–886.
- Bellamy, C., Boughey, K., Hawkins, C., Reveley, S., Spake, R., Williams, C. and Altringham, J. 2020. A sequential multi-level framework to improve habitat suitability modelling. – *Landscape Ecol.* 35: 1001–1020.
- Booth, T. H. 2018. Species distribution modelling tools and databases to assist managing forests under climate change. – *For. Ecol. Manage.* 430: 196–203.
- Breiman, L. 2001. Random forests. – *Mach. Learn.* 45: 5–32.
- 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. and Guisan, A. 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. – *Global Ecol. Biogeogr.* 21: 481–497.
- Carroll, C., Parks, S. A., Dobrowski, S. Z. and Roberts, D. R. 2018. Climatic, topographic, and anthropogenic factors determine connectivity between current and future climate analogs in North America. – *Global Change Biol.* 24: 5318–5331.
- Charney, N. D., Record, S., Gerstner, B. E., Merow, C., Zarnetske, P. L. and Enquist, B. J. 2021. A test of species distribution model transferability across environmental and geographic space for 108 western North American tree species. – *Front. Ecol. Evol.* 9: 689295.
- Chevalier, M., Broennimann, O., Cornuault, J. and Guisan, A. 2021. Data integration methods to account for spatial niche

- truncation effects in regional projections of species distribution. – *Ecol. Appl.* 31: e02427.
- Chevalier, M., Zarzo-Arias, A., Guélat, J., Mateo, R. G. and Guisan, A. 2022. Accounting for niche truncation to improve spatial and temporal predictions of species distributions. – *Front. Ecol. Evol.* 10: 944116.
- Di Cola, V., Broennimann, O., Petitpierre, B., Randin, C., Engler, R., Dubuis, A., D'Amen, M., Pellissier, L., Pottier, J., Pio, D., Salamin, N., Breiner, F. T., Mateo, R. G., Hordijk, W. and Guisan, A. 2017. Ecospat: an R package for the support of spatial analyses and modelling of species niches and distributions. – *Ecography* 40: 774–787.
- Dubuis, A., Pottier, J., Rion, V., Pellissier, L., Theurillat, J. P. and Guisan, A. 2011. Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. – *Divers. Distrib.* 17: 1122–1131.
- Dyderski, M. K., Paż, S., Frelich, L. E. and Jagodziński, A. M. 2017. How much does climate change threaten European forest tree species distributions? – *Global Change Biol.* 24: 1150–1163.
- Elith, J., Kearney, M. and Phillips, S. 2010. The art of modelling range-shifting species. – *Methods Ecol. Evol.* 1: 330–342.
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E. and Yates, C. J. 2011. A statistical explanation of MaxEnt for ecologists. – *Divers. Distrib.* 17: 43–57.
- Fan, J., Upadhye, S. and Worster, A. 2006. Understanding receiver operating characteristic (ROC) curves. – *Can. J. Emerg. Med.* 8: 19–20.
- Gallien, L., Douzet, R., Pratte, S., Zimmermann, N. E. and Thuiller, W. 2012. Invasive species distribution models - how violating the equilibrium assumption can create new insights. – *Global Ecol. Biogeogr.* 21: 1126–1136.
- Gastón, A. and García-Viñas, J. I. 2010. Updating coarse-scale species distribution models using small fine-scale samples. – *Ecol. Modell.* 221: 2576–2581.
- Gastón, A., Ciudad, C., Mateo-Sánchez, M. C., García-Viñas, J. I., López-Leiva, C., Fernández-Landa, A., Marchamalo, M., Cuevas, J., de la Fuente, B., Fortin, M. J. and Saura, S. 2017. Species' habitat use inferred from environmental variables at multiple scales: how much we gain from high-resolution vegetation data? – *Int. J. Appl. Earth Obs. Geoinf.* 55: 1–8.
- Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – *Ecol. Modell.* 135: 147–186.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – *Ecol. Lett.* 8: 993–1009.
- Guisan, A., Edwards Jr, T. C. and Hastie, T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. – *Ecol. Modell.* 157: 89–100.
- Guisan, A. et al. 2013. Predicting species distributions for conservation decisions. – *Ecol. Lett.* 16: 1424–1435.
- Guisan, A., Thuiller, W. and Zimmermann, N. 2017. Habitat suitability and distribution models: with applications in R. – Cambridge Univ. Press.
- Hannemann, H., Willis, K. J. and Macias-Fauria, M. 2016. The devil is in the detail: unstable response functions in species distribution models challenge bulk ensemble modelling. – *Global Ecol. Biogeogr.* 25: 26–35.
- Hastie, T. and Tibshirani, R. 1986. Generalized additive models. – *Stat. Sci.* 1: 297–318.
- Hickling, R., Roy, D. B., Hill, J. K., Fox, R. and Thomas, C. D. 2006. The distributions of a wide range of taxonomic groups are expanding polewards. – *Global Change Biol.* 12: 450–455.
- Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C. and Guisan, A. 2006. Evaluating the ability of habitat suitability models to predict species presences. – *Ecol. Modell.* 199: 142–152.
- Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P. and Kessler, M. 2017. Climatologies at high resolution for the earth's land surface areas. – *Sci. Data* 4: 170122.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. and Liu, T.-Y. 2017. LightGBM: an effective decision tree. – *Adv. Neural Inf. Process. Syst.* 30: 1111–1116.
- Keil, P., Belmaker, J., Wilson, A. M., Unitt, P. and Jetz, W. 2013. Downscaling of species distribution models: a hierarchical approach. – *Methods Ecol. Evol.* 4: 82–94.
- Kéry, M. and Royle, J. A. 2010. Hierarchical modelling and estimation of abundance and population trends in metapopulation designs. – *J. Anim. Ecol.* 79: 453–461.
- Kindt, R. 2023. TreeGOER: a database with globally observed environmental ranges for 48,129 tree species. – *Global Change Biol.* 29: 6303–6318.
- Leroy, B., Meynard, C. N., Bellard, C. and Courchamp, F. 2016. virtualspecies, an R package to generate virtual species distributions. – *Ecography* 39: 599–607.
- Loarie, S. R., Duffy, P. B., Hamilton, H., Asner, G. P., Field, C. B. and Ackerly, D. D. 2009. The velocity of climate change. – *Nature* 462: 1052–1055.
- Lomba, A., Pellissier, L., Randin, C., Vicente, J., Moreira, F., Honrado, J. and Guisan, A. 2010. Overcoming the rare species modelling paradox: a novel hierarchical framework applied to an Iberian endemic plant. – *Biol. Conserv.* 143: 2647–2657.
- Mackey, B. G. and Lindenmayer, D. B. 2001. Towards a hierarchical framework for modelling the spatial distribution of animals. – *J. Biogeogr.* 28: 1147–1166.
- Maiorano, L., Cheddadi, R., Zimmermann, N. E., Pellissier, L., Petitpierre, B., Pottier, J., Laborde, H., Hurdu, B. I., Pearman, P. B., Psomas, A., Singarayer, J. S., Broennimann, O., Vittoz, P., Dubuis, A., Edwards, M. E., Binney, H. A. and Guisan, A. 2013. Building the niche through time: using 13,000 years of data to predict the effects of climate change on three tree species in Europe. – *Global Ecol. Biogeogr.* 22: 302–317.
- Márcia Barbosa, A., Real, R., Muñoz, A. R. and Brown, J. A. 2013. New measures for assessing model equilibrium and prediction mismatch in species distribution models. – *Divers. Distrib.* 19: 1333–1338.
- Mateo, R. G., Mokany, K. and Guisan, A. 2017. Biodiversity models: what if unsaturation is the rule? – *Trends Ecol. Evol.* 32: 556–566.
- Mateo, R. G., Aroca-Fernández, M. J., Gastón, A., Gómez-Rubio, V., Saura, S. and García-Viñas, J. I. 2019a. Looking for an optimal hierarchical approach for ecologically meaningful niche modelling. – *Ecol. Modell.* 409: 108735.
- Mateo, R. G., Gastón, A., Aroca-Fernández, M. J., Broennimann, O., Guisan, A., Saura, S. and García-Viñas, J. I. 2019b. Hierarchical species distribution models in support of vegetation conservation at the landscape scale. – *J. Veg. Sci.* 30: 386–396.
- Mauri, A., Strona, G. and San-Miguel-Ayanz, J. 2017. EU-Forest, a high-resolution tree occurrence dataset for Europe. – *Sci. Data* 4: 160123.
- McCullagh, P. and Nelder, J. A. 1989. Generalized linear models. – Chapman & Hall.
- McGarigal, K., Wan, H. Y., Zeller, K. A., Timm, B. C. and Cushman, S. A. 2016. Multi-scale habitat selection modeling: a review and outlook. – *Landscape Ecol.* 31: 1161–1175.
- McGill, B. J. 2010. Matters of scale. – *Science* 328: 575–576.

- Mod, H. K., Scherrer, D., Luoto, M. and Guisan, A. 2016. What we use is not what we know: environmental predictors in plant distribution models. – *J. Veg. Sci.* 27: 1308–1322.
- Moreno-Amat, E., Mateo, R. G., Nieto-Lugilde, D., Morueta-Holme, N., Svenning, J. C. and García-Amorena, I. 2015. Impact of model complexity on cross-temporal transferability in Maxent species distribution models: an assessment using paleobotanical data. – *Ecol. Modell.* 312: 308–317.
- Moudry, V. et al. 2023. Optimising species distribution models: sample size, positional error, and sampling bias matter. – *Eco-evoRxiv* 1–28.
- Normand, S., Ricklefs, R. E., Skov, F., Bladt, J., Tackenberg, O. and Svenning, J. C. 2011. Postglacial migration supplements climate in determining plant species ranges in Europe. – *Proc. R. Soc. B* 278: 3644–3653.
- Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P. and Kassem, K. R. 2001. Terrestrial ecoregions of the world: a new map of life on Earth. – *BioScience* 51: 933–938.
- Owens, H. L., Campbell, L. P., Dornak, L. L., Saupe, E. E., Barve, N., Soberón, J., Ingenloff, K., Lira-Noriega, A., Hensz, C. M., Myers, C. E. and Peterson, A. T. 2013. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. – *Ecol. Modell.* 263: 10–18.
- Pearman, P. B., Randin, C. F., Broennimann, O., Vitzthum, P., Knaap, W. O. V. Der, Engler, R., Lay, G. Le, Zimmermann, N. E. and Guisan, A. 2008. Prediction of plant species distributions across six millennia. – *Ecol. Lett.* 11: 357–369.
- Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – *Global Ecol. Biogeogr.* 12: 361–371.
- Pearson, R. G., Dawson, T. P. and Liu, C. 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. – *Ecography* 27: 285–298.
- Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. and Guisan, A. 2012. Climatic niche shifts are rare among terrestrial plant invaders. – *Science* 335: 1344–1348.
- Petitpierre, B., McDougall, K., Seipel, T., Broennimann, O., Guisan, A. and Kueffer, C. 2016. Will climate change increase the risk of plant invasions into mountains? – *Ecol. Appl.* 26: 530–544.
- Phillips, S. J., Anderson, R. P. and Schapire, R. E. 2006. Maximum entropy modeling of species geographic distributions. – *Ecol. Modell.* 190: 231–259.
- Qiao, H., Feng, X., Escobar, L. E., Peterson, A. T., Soberón, J., Zhu, G. and Papeş, M. 2019. An evaluation of transferability of ecological niche models. – *Ecography* 42: 521–534.
- Rousseau, J. S. and Betts, M. G. 2022. Factors influencing transferability in species distribution models. – *Ecography* 2022: e06060.
- Scherrer, D., Esperon-Rodriguez, M., Beaumont, L. J., Barradas, V. L. and Guisan, A. 2021. National assessments of species vulnerability to climate change strongly depend on selected data sources. – *Divers. Distrib.* 27: 1367–1382.
- Serra-Varela, M. J., Grivet, D., Vincenot, L., Broennimann, O., Gonzalo-Jiménez, J. and Zimmermann, N. E. 2015. Does phylogeographical structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.). – *Global Ecol. Biogeogr.* 24: 1302–1313.
- Sofaer, H. R., Jarnevich, C. S., Pearse, I. S., Smyth, R. L., Auer, S., Cook, G. L., Edwards, T. C., Guala, G. F., Howard, T. G., Morisette, J. T. and Hamilton, H. 2019. Development and delivery of species distribution models to inform decision-making. – *BioScience* 69: 544–557.
- Tang, W., Hu, J., Zhang, H., Wu, P. and He, H. 2015. Kappa coefficient: a popular measure of rater agreement. – *Shanghai Arch. Psychiatry* 27: 62–67.
- Thuiller, W., Brotons, L., Araújo, M. B. and Lavorel, S. 2004. Effects of restricting environmental range of data to project current and future species distributions. – *Ecography* 27: 165–172.
- Titeux, N., Maes, D., Van Daele, T., Onkelinx, T., Heikkinen, R. K., Romo, H., García-Barros, E., Munguira, M. L., Thuiller, W., van Swaay, C. A. M., Schweiger, O., Settele, J., Harpke, A., Wiemers, M., Brotons, L. and Luoto, M. 2017. The need for large-scale distribution data to estimate regional changes in species richness under future climate change. – *Divers. Distrib.* 23: 1393–1407.
- Tulloch, A. I. T., Sutcliffe, P., Naujokaitis-Lewis, I., Tingley, R., Brotons, L., Ferraz, K. M. P. M. B., Possingham, H., Guisan, A. and Rhodes, J. R. 2016. Conservation planners tend to ignore improved accuracy of modelled species distributions to focus on multiple threats and ecological processes. – *Biol. Conserv.* 199: 157–171.
- Valavi, R., Elith, J., Lahoz-Monfort, J. J. and Guillerá-Arroita, G. 2021. Modelling species presence-only data with random forests. – *Ecography* 44: 1731–1742.
- Vicente, J. R., Gonçalves, J., Honrado, J. P., Randin, C. F., Pottier, J., Broennimann, O., Lomba, A. and Guisan, A. 2014. A framework for assessing the scale of influence of environmental factors on ecological patterns. – *Ecol. Complex* 20: 151–156.
- Webber, B. L., Yates, C. J., Le Maitre, D. C., Scott, J. K., Kriticos, D. J., Ota, N., McNeill, A., Le Roux, J. J. and Midgley, G. F. 2011. Modelling horses for novel climate courses: insights from projecting potential distributions of native and alien Australian acacias with correlative and mechanistic models. – *Divers. Distrib.* 17: 978–1000.
- Williams, J. W. and Jackson, S. T. 2007. Novel climates, no-analog communities, and ecological surprises. – *Front. Ecol. Environ.* 5: 475–482.
- Yates, K. L. et al. 2018. Outstanding challenges in the transferability of ecological models. – *Trends Ecol. Evol.* 33: 790–802.
- Zurell, D., Berger, U., Cabral, J. S., Jeltsch, F., Meynard, C. N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B. and Grimm, V. 2010. The virtual ecologist approach: simulating data and observers. – *Oikos* 119: 622–635.
- Zurell, D., Elith, J. and Schröder, B. 2012. Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions. – *Divers. Distrib.* 18: 628–634.
- Zurell, D. et al. 2020. A standard protocol for reporting species distribution models. – *Ecography* 43: 1261–1277.