www.nature.com/scientificdata

scientific data

Check for updates

OPEN

DATA DESCRIPTOR

SDMapCH: a Comprehensive database of >7,500 modelled species habitat suitability maps for Switzerland

Antoine Adde[1]✉, Pierre-Louis Rey[2], Nathan Külling[3], Yohann Chauvier-Mendes[1,4], Fabian Fopp[4,5], Manuel R. Popp[4], Olivier Broennimann[2,6], Blaise Petitpierre[7], Nicolas Strebel[8], Andrin Gross[9], Silvia Stofer[10], Anthony Lehmann[3], Niklaus E. Zimmermann[4], Loïc Pellissier[4,5], Antoine Guisan[2,6]✉ & Florian Altermatt[1,11]✉

Conserving natural ecosystems requires consistent and standardized biodiversity data to advance scientific research and ecological understanding. Despite several national initiatives to develop databases of species habitat suitability maps, even well-studied countries often lack comprehensive, standardized databases that cover a wide range of taxonomic groups modelled using a consistent framework. Using Switzerland as a case study, we demonstrate how these gaps can be addressed by introducing SDMapCH (v1.3), a nationwide raster database of species habitat suitability maps at 25-meter resolution. SDMapCH provides maps for about 7,500 species under both present conditions and future climate scenarios. SDMapCH was developed using the N-SDM software, an end-to-end platform based on a spatially-nested hierarchical framework. N-SDM allows multi-level integration of species and covariate data, helping to address niche truncation. SDMapCH outputs were evaluated using a state-of-the-art cross-validation procedure, and all layers passed a systematic data integrity check. By providing standardized, high-resolution habitat suitability maps for diverse species across various taxonomic and functional groups, SDMapCH stands as a key resource for scientific research and biodiversity assessments.

Background & Summary

Species Distribution Models (SDMs; also commonly referred to as Habitat Suitability Models HSMs) have become essential tools in biodiversity and conservation science, generalizing information on species' habitat suitability across geographic space and providing insights into the environmental factors at play[1–3]. Habitat suitability maps derived from SDMs have proven valuable resources across multiple domains of ecological research and conservation applications, such as ecological infrastructure planning, habitat restoration, invasive species management, and climate change adaptation[4–6].

[1]Department of Aquatic Ecology, Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland. [2]Institute of Earth Surface Dynamics, Faculty of Geosciences and Environment, University of Lausanne, Lausanne, Switzerland. [3]EnviroSPACE, Institute for Environmental Sciences, University of Geneva, Geneva, Switzerland. [4]Land Change Science Research Unit, Swiss Federal Institute for Forest, Snow and Landscape Research, WSL, Birmensdorf, Switzerland. [5]Ecosystems Landscape Evolution, Institute for Terrestrial Ecosystems, Department of Environmental System Sciences, ETH Zurich, Zurich, Switzerland. [6]Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland. [7]InfoFlora, Chambésy-Geneva, Switzerland. [8]Swiss Ornithological Institute, Sempach, Switzerland. [9]National data and information centre for the documentation and conservation of the Swiss fungi flora (SwissFungi), Swiss Federal Institute for Forest, Snow and Landscape Research, WSL, Birmensdorf, Switzerland. [10]National data and information centre of Swiss lichens (SwissLichens), Swiss Federal Institute for Forest, Snow and Landscape Research, WSL, Birmensdorf, Switzerland. [11]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland. ✉e-mail: antoine.adde@eawag.ch; antoine.guisan@unil.ch; florian.altermatt@ewag.ch

Standardized habitat suitability maps, developed using state-of-the-art methods and covering an extensive range of species, are essential for informed biodiversity assessments[6–8]. In response, several countries have launched initiatives to create comprehensive and standardized national databases of species' habitat suitability maps. Examples of such national-level initiatives include the effort by NatureServe to map habitat suitability for 2,216 at-risk species across the contiguous United States (https://www.natureserve.org/), and a nationwide database in Australia covering 1,872 terrestrial and freshwater vertebrate species[9]. At a continental scale, the European Breeding Bird Atlas provides modelled habitat suitability maps for over 500 bird species across Europe[10]. However, even in well-studied countries in Western Europe and North America, these efforts often focus on a few well-known taxonomic groups, while many others remain underrepresented or entirely omitted. Switzerland is a typical example of this issue: despite considerable efforts by the Swiss Species Information Centers (InfoSpecies; www.infospecies.ch) to compile extensive species records and display maps based on observational data, many taxonomic groups remain only partially mapped or entirely unmapped.

By generalizing information from observational data, even when sparse, modelled habitat suitability maps provide a crucial complement. Moreover, when generated using a standardized and unified method, these maps enable systematic comparisons across species groups. Here, we demonstrate how such upscaling and generalization can be achieved at the national scale for thousands of species, using Switzerland as a case example. Specifically, we introduce SDMapCH (v1.3), a nationwide raster database of modelled species habitat suitability maps at a 25-meter resolution. SDMapCH provides individual maps for about 7,500 species, as well as aggregated maps for 11 major taxonomic groups, 26 ecological guilds, and 5 conservation status categories, under both present conditions and future climate change scenarios projected through the end of the century. SDMapCH was developed using the N-SDM modelling pipeline[11], an end-to-end platform built on a spatially nested hierarchical framework and optimized for high-performance computing environments. N-SDM enables multi-level and multi-resolution integration of species and covariate data, addressing niche truncation issues[12], and supporting parallel high-resolution modelling of numerous species within tight timeframes.

By providing standardized, high-resolution habitat suitability maps for diverse species across various taxonomic and functional groups, SDMapCH stands as a key resource for advancing scientific analyses and streamlining workflows for a wide range of applications in Switzerland and beyond. These applications include, but are not limited to, biological conservation in changing ecosystems[13–15], biodiversity-energy development conflicts[16–18], and evaluating nature's contributions to people[19–21]. Since species, environmental data, and modelling techniques are constantly evolving, SDMapCH was structured to allow updates through the release of new versions. The technical characteristics and parameters specified in this article are those used for the first open version of SDMapCH (v1.3), following previous preliminary versions. The SDMapCH database and associated metadata, including the ODMAP reporting protocol[22], detailed lists of species and candidate covariates, information on covariate selection and importance, and model evaluation results, are openly available on Dryad.

Methods

Species' habitat suitability modelling. To model and map species' habitat suitability, we used the N-SDM software[11]. N-SDM is an end-to-end SDM platform built around a spatially-nested framework. It combines a 'global' model, which quantifies the species' response to bioclimatic conditions across its entire distributional range, with a 'regional' model fitted using finer-scale habitat covariates. Global and regional models were combined by using the 'covariate' nesting strategy, which consists in using the global model predictions (i.e., bioclimatic envelope) as a forced additional covariate for fitting the regional model. In the metadata.zip file accompanying the maps on Dryad, we provided both the N-SDM configuration file containing user-specific settings for all editable options, and the ODMAP reporting protocol[22].

Study areas. Following the spatially-nested framework of N-SDM, we distinguished between regional- and global-level study areas. The regional-level area (projection extent) included all of Switzerland, with a total area of about 41,000 km². For the global-level area (calibration extent), we used a bounding box approximately 250 times larger, covering the European continent and a portion of Northern Africa (ranging from 32.60 °N to 71.70 °N, and from 28.56 °W to 40.21 °E). The projection and calibration extents were the same for all species, selected as a tradeoff to encompass most species' climatic niches while keeping computational costs manageable. Moreover, incorporating data from distant continents might increase uncertainty due to data consistency issues and taxonomic differences[23–25]. This extent also balanced spatial scale and model accuracy[26–28], preventing niche truncation for widely distributed species and maintaining accuracy for those limited to specific locations.

Species data. Two sets of species occurrence records were used: one for the regional level (within Switzerland) and one for the global level (outside Switzerland). For Switzerland, validated occurrence records for 21,290 species aggregated at a 25-m resolution and covering the 1980–2021 period were provided by the Swiss Species Information Center InfoSpecies (https://doi.org/10.15468/x9z635)[29]. To avoid modelling issues related to low sample size (e.g., model convergence, number of covariates), only species with more than 50 occurrence records after spatial thinning (described below) were retained for modelling, resulting in 7,508 species (species list with detailed taxonomic information available in the metadada.zip file accompanying the maps on Dryad). Species were classified into 11 major taxonomic groups: amphibians (16 species modelled out of 21 known to be part of Switzerland's native biodiversity[30]), arthropods (2,266/32,685; including 296 Ephemeroptera, Plecoptera, Trichoptera, and Odonata (EPTO)), breeding birds (175/205), ferns and mosses (589/1,236), fishes (30/71), fungi and lichens (1,700/11,121), mammals (77/86), molluscs (157/270), reptiles (14/16), and vascular plants (2,096/2,536). The 378 remaining modelled species were nonnative to Switzerland, including 314 vascular plants, 19 arthropods, and 13 fungi and lichens. In addition, when available, InfoSpecies provided information on the ecological guild of each species, totaling 26 guilds (i.e., groups of species with similar ecological needs, based on the typology of Swiss natural habitats[31]). Red List conservation status for each species in Switzerland[30] were also provided by InfoSpecies. Occurrence records for outside of Switzerland for matching species and the same 1980–2021 period were obtained from the Global Biodiversity Information Facility (GBIF; https://www.gbif.org/) (https://doi.org/10.15468/dl.fktyas). The extracted GBIF records were tested and filtered for coordinate validity, which included identifying equal latitude/longitude values, rounded coordinates, and zero values[32]. Additionally, we filtered for spatial and temporal outliers by extracting occurrences only within the area defined by a shapefile covering the target European domain and specifying the same period as the Swiss data (1980–2021). Further details on GBIF data processing and matching with InfoSpecies species names are provided in the metadada.zip file accompanying the maps on Dryad. Sampling bias was corrected using a spatial filter approach that thinned the occurrence records, mitigating observation clusters[33–35]. We applied level-specific minimum distances between two observations for spatial thinning: 500 m for the regional level and 1 km for the global level. Ideally, these values should be based on species biology, environmental conditions, and previous research on the study system[34,36,37]. However, such comprehensive information was not available for all species. Therefore, we opted for a tradeoff, considering the spatial resolution of the covariate data (25 m within Switzerland and 30 arcseconds outside) and choosing conservatively larger distances to ensure efficient spatial disaggregation. Finally, for each species and occurrence set, 10,000 background points were randomly generated, using a single background approach for efficiency since differences between methods are usually minimal[38–40].

Covariate data. To model the habitat suitability of each individual species, we used a suite of 877 candidate covariates from eight main categories (bioclimatic, edaphic, hydrologic, land use and cover, population density, topographic, transportation, and vegetation; detailed covariate list available in the metadada.zip file accompanying the maps on Dryad). Only bioclimatic covariates ($n = 19$) were used for global-level models, while all other covariates were used for regional-level models. Regional-level covariates were extracted from the 25-m resolution SWECO25 raster database[41]. To capture environmental conditions across a broader area surrounding the occurrence records, habitat covariates were retrieved using 13 focal moving windows available in SWECO25, each computing the average pixel value within circular neighborhoods of increasing size, with radii ranging from 25 m to 5 km. For global-level covariates, to benefit from best-available products, we used specific data sources for inside and outside of Switzerland. For Switzerland, bioclimatic covariates were extracted from the CHclim25 dataset (https://doi.org/10.5281/zenodo.10635681), which gather Swiss-wide climate layers downscaled at 25-m resolution. For outside of Switzerland, we used 30-arcsecond CHELSA layers[42]. To project potential future habitat suitability, we retrieved projection data of all 19 bioclimatic covariates from the CHclim25 dataset for the four coupled global and regional climate models (GCM-RCM) available (CLMCOM-CCLM4/HADGEM, DMI-HIRHAM/ECEARTH, MPICSC-REMO2/MPIESM, and SMHI-RCA/IPSL), three future intervals (2020–2049, 2045–2074, and 2070–2099), and two greenhouse gas concentration trajectories or representative concentration pathways (RCPs)[43], RCP4.5 ("Low Carbon") and RCP8.5 ("High Carbon"). All covariates were standardized to zero mean and unit variance.

Covariate selection. As a preliminary step, candidate covariates identified by experts as being likely irrelevant for modelling a species (or group of species) were discarded by providing N-SDM with a so called 'expert table' (available in the metadata.zip file accompanying the maps on Dryad) filled in collaboration with species experts. For instance, edaphic covariates, such as soil pH or aeration, were not used for modelling butterflies and reptile species, and hydrological covariates related to stream morphology were discarded for modelling fungi or bat species. By doing so, we ensured that only biologically realistic and relevant variables were considered. The two-step automated covariate selection procedure included in N-SDM, 'covsel'[44], was then applied for each species at both global and regional levels, with default values. In short, covsel first reduces the dimensionality of the candidate set by eliminating less informative covariates among collinear pairs (default Pearson correlation coefficient: $|r| > 0.70$) and remaining covariates are used to fit models with algorithm-specific embedded regularization procedures. Covariates are ranked on the basis of their importance scores and the top k covariates are retained, with default $k = \text{round}(\log2(\text{number of occurrences}))$. To avoid models ending with too many covariates, the default maximum possible value for k is 12. Since the covariate nesting strategy was used, the global-model output was forced as an additional covariate in all regional models. Relative covariate importance, ranging between 0 and 1, was automatically outputted by N-SDM (see help files of the 'nsdm.varimp' function for more details). Finally, in an attempt to mitigate potential observational biases in occurrence data, we used the N-SDM option that allows setting a zero value to the set of selected covariates expected to be related to these biases, including distance from the transportation network or settlements (see the N-SDM configuration file available in the metadada.zip file accompanying the maps on Dryad for the full list of observational covariates). This approach can help reduce potential observer bias in the prediction area, provided that the selected covariates effectively capture the spatial patterns of observer activity[45–47].

Model fitting and selection. The five modelling algorithms available in N-SDM – Generalized Linear Model (GLM)[48], Generalized Additive Model (GAM)[49], Maxnet (MAX)[50], Random Forest (RF)[51], and light Gradient Boosted Machine (GBM)[52] – were fitted using their default N-SDM values for hyperparameter tuning. To account for class imbalance, background points and occurrences were weighted equally in the models[39]. Model accuracy was evaluated using a split-sample approach repeated 100 times with 30% of the data kept for validation[2]. For each model, the best combination of hyperparameters was identified using the average "Score"[11] of three evaluation metrics including the re-scaled Area Under the Curve′ (AUC′; or Somers' D, such as $AUC' = AUC * 2 - 1$)[53], the maximized True Skill Statistic (maxTSS)[2], and the Continuous Boyce Index (CBI)[54]. All these evaluation metrics ranged between -1 and $+1$.

Spatial predictions. We predicted habitat suitability values for each species and five modelling algorithms for the baseline period and three future intervals × two RCPs. Results from the five modelling algorithms were individually mapped over a 25-m resolution grid of 64,007,390 cells covering Switzerland (CH1903+/ LV95 geographic coordinate system), and ensembled together by computing the mean of the five maps. Aggregated maps for the 11 major taxonomic groups, 26 ecological guilds, and 5 conservation status categories were computed by taking the mean of the maps corresponding to the subset of species within each group. Maps were provided for both mean values and the coefficient of variation.

Data Records

The SDMapCH (v1.3) database is publicly available on Dryad[55]. It consists of five main map archives (tar.zst) and one companion metadata archive (.zip). SDMapCHv1_3_current.tar.zst contains GeoTIFF (.tif) maps for the current period. SDMapCHv1_3_2020_2049.tar.zst, SDMapCHv1_3_2045_2074.tar.zst, and SDMapCHv1_3_2070_2099.tar.zst contain maps for the three future periods (2020–2049, 2045–2074, 2070–2099) under the two RCP scenarios (RCP4.5 and RCP8.5). SDMapCHv1_3_aggregates.tar.zst contains aggregated maps for species grouped by ecological guild, Swiss Red List status, major taxonomic group, and all species combined, for the current and three future periods. Finally, metadata.zip contains supporting tabular and text information in CSV and PDF format.

Data Overview

Individual habitat suitability maps were compiled in SDMapCH (v1.3)[55] for 7,425 species out of 7,508 modelled, excluding 83 sensitive species (1.1%) for which the dissemination of distribution data posed conservation or legal concerns. These concerns relate to the risk of collection, destruction, or disturbance of rare and endangered species, such as certain orchids and reptiles, their potential exploitation through the international illegal wildlife trade, and the requirements of Swiss biodiversity conservation law. However, individual habitat suitability maps for all these sensitive species are available upon formal request to InfoSpecies (Swiss Species Information Centre) by contacting info.species@unine.ch, in accordance with their standard access procedure. In addition to the individual maps, aggregated versions (mean and coefficient of variation) for the 11 major taxonomic groups, 26 ecological guilds, and 5 conservation status categories were included, resulting in a total size of 487 GB. We illustrated with maps obtained for individual species (Fig. 1) and major taxonomic groups (Fig. 2) for the current period and future projections.

Technical Validation

Model evaluation. Figure 3a displays cross-validated evaluation metrics obtained for the final ensemble maps of the 7,508 individual species. The mean ± standard deviation Score value was 0.89 ± 0.05 indicating rather high predictive performances. We provided detailed model evaluation metrics in the metadata.zip file accompanying the maps on Dryad, including breakdowns of AUC′, maxTSS, CBI, and their average Score, summarized by taxonomic group and modelling algorithm, including the ensemble version. Across all metrics, GLM showed the lowest mean values and the highest standard deviations. MAX, RF, GBM, and RF achieved the highest mean values for AUC′, CBI, maxTSS, and Score, respectively. Group-wise, amphibians achieved the highest ensemble performance for AUC′, maxTSS, and Score, whereas birds had the highest CBI. Fungi and lichens tended to have the lowest ensemble scores across all metrics. For each group, we compared the ensemble score with the highest score obtained by any single algorithm. On average at the group level, ensembles performed close to the best single algorithm, with differences ranging from −0.001 to −0.009 for AUC′, −0.012 to −0.025 for maxTSS, −0.041 to −0.111 for CBI, and −0.017 to −0.034 for Score. These results indicate that ensembles deliver consistently high and stable performance across taxa, although in all cases a single algorithm slightly outperformed them.

In addition to quantitative evaluation metrics, taxonomic experts from InfoSpecies centers qualitatively assessed the plausibility of habitat suitability maps for around 100 selected species across various taxonomic groups (e.g., birds, plants, reptiles). Overall, expert feedback highlighted that the models from SDMapCH v1.3 presented in this article were well-calibrated and performed better than previous preliminary versions (v1.1 and v1.2). As the modelling approach was not optimized for individual species or species groups, some models exhibited effects of sampling biases present in the original datasets, resulting in localized variations in predictions. For example, certain habitats (e.g., wetlands vs. forests) may be over- or under-represented in occurrence records for specific groups, leading to corresponding over- or under-predictions in habitat suitability[56]. More specifically, several reptile models showed underrepresentation of favorable areas in southwestern and southeastern parts of the country, with certain species like *Natrix helvetica* predicted too narrowly despite a known broader range. Bird models were identified by the experts as showing slight signals of detection bias in the raw observation data, with species generally being more elusive in their natural habitats (e.g., *Accipiter gentilis*, *Accipiter nisus*) being underpredicted, while birds that are much more commonly reported, such as from urban areas with higher observation effort (e.g., *Acrocephalus palustris*) being slightly overpredicted; in addition, probably due to missing key habitat covariates, reduced accuracy was expertized for wetland specialists (e.g., *Acrocephalus scirpaceus*, *Actitis hypoleucos*). Butterfly models were generally well-calibrated but exhibited localized overprediction for some species, such as *Coenonympha glycerion* in southwestern regions, and in some cases reflected biases linked to agricultural and garden habitats (e.g., *Pieris napi*). Plant models sometimes lacked environmental covariates for edaphic or alluvial specialists (e.g., *Myricaria germanica*, *Heracleum mantegazzianum*), leading to restrictive predictions for certain species (e.g., *Prunus serotina*, *Vitis sylvestris*) and overprediction in specific areas (e.g., *Agrostis canina* in southwestern regions). Mammal models were generally robust, but localized overprediction were reported for species such as *Crocidura leucodon* and *Felis silvestris*, particularly in southwestern regions. While refinements in data processing, covariate selection, and species-specific optimization could improve the ecological robustness of future versions, such efforts are beyond the scope of this

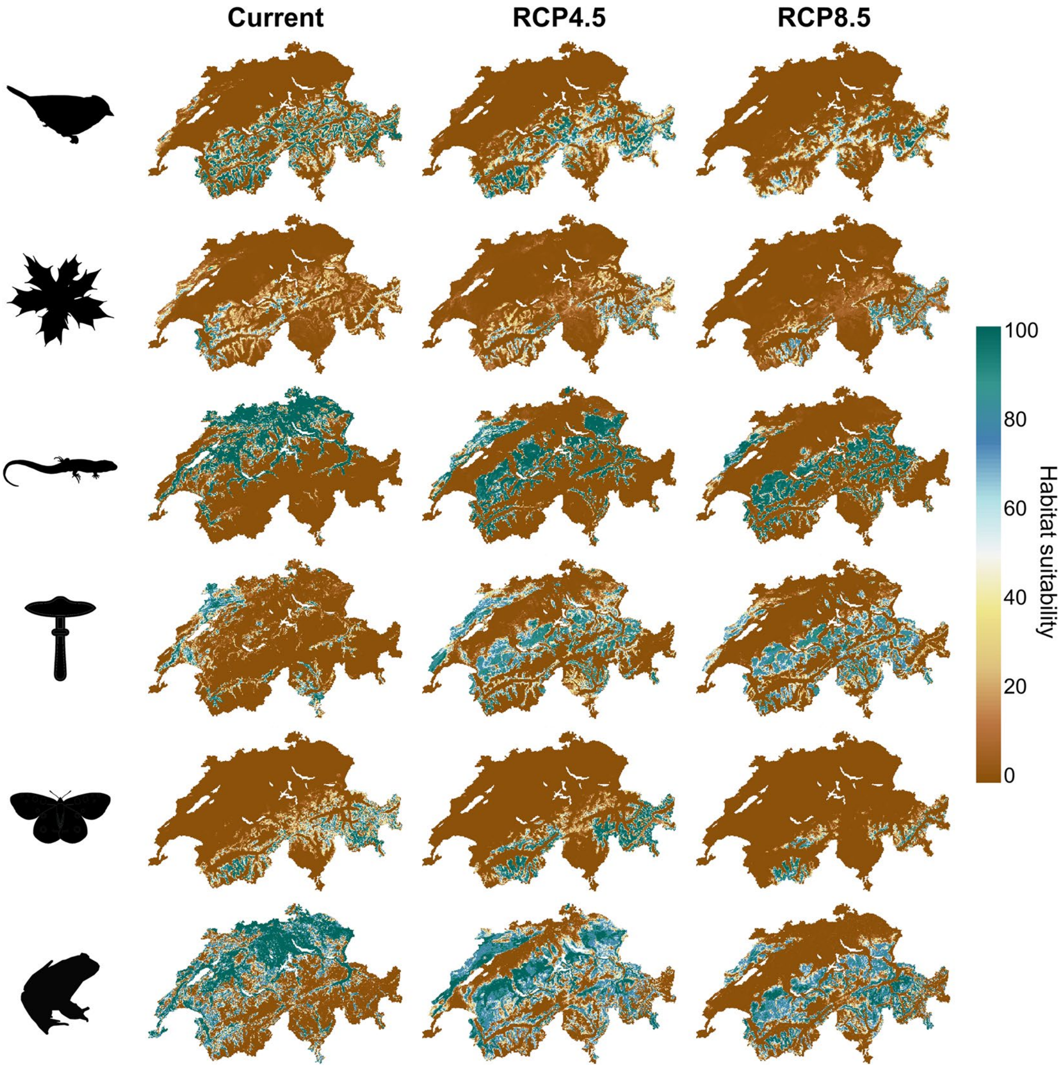

Fig. 1 SDMapCH (v1.3) habitat suitability maps for example species for current and future (2070–2099) periods under the low (RCP4.5) and high (RCP8.5) greenhouse gas scenarios. For each species, values were rescaled to a 0–100 range for data storage purposes and improved readability, with, from top to bottom: *Acanthis flammea* (bird), *Eryngium alpinum* (vascular plant), *Lacerta agilis* (reptile), *Macrolepiota procera* (fungus), *Parnassius phoebus* (arthropod), and *Rana temporaria* (amphibian). For the use and validation of individual habitat suitability maps, please note: (i) The habitat suitability maps were primarily developed for research purposes; (ii) These maps were not validated for each individual species; therefore, interpretation at the species level should be complemented with additional taxonomic or faunistic expertise; (iii) The results of these models were not systematically validated by InfoSpecies; (iv) Individual maps should not be used directly for practical applications or management decisions without consultation with experts in the relevant taxonomic groups. The authors disclaim responsibility for the accuracy of individual maps.

work, which focuses on applying a consistent and coherent approach across all species, and not taxon-specific optimization. While such an individual optimization would offer the benefit of more finely tuned maps for each species, it comes at the cost of heterogeneous modelling approaches that hinder comparability. Therefore, the use of a harmonized methodology across over 7,500 species represents a key strength for broad-scale biodiversity analyses. For species-specific conservation planning, however, tailored refinements should be applied, ideally in collaboration with relevant taxonomic and faunistic experts.

Covariate selection and importance. All 19 candidate bioclimatic covariates at the global level were included in at least one model. BIO15 (precipitation seasonality), BIO1 (annual mean temperature), and BIO2 (mean diurnal range) were the three most frequently selected global-level covariates, by 7,452, 7,431, and 7,271 models, respectively. On average, BIO1 (annual mean temperature) was the global-level covariate with the highest relative importance (0.86 ± 0.13), followed by BIO4 (temperature seasonality; 0.73 ± 0.15), and BIO10 (mean temperature of warmest quarter; 0.65 ± 0.17).

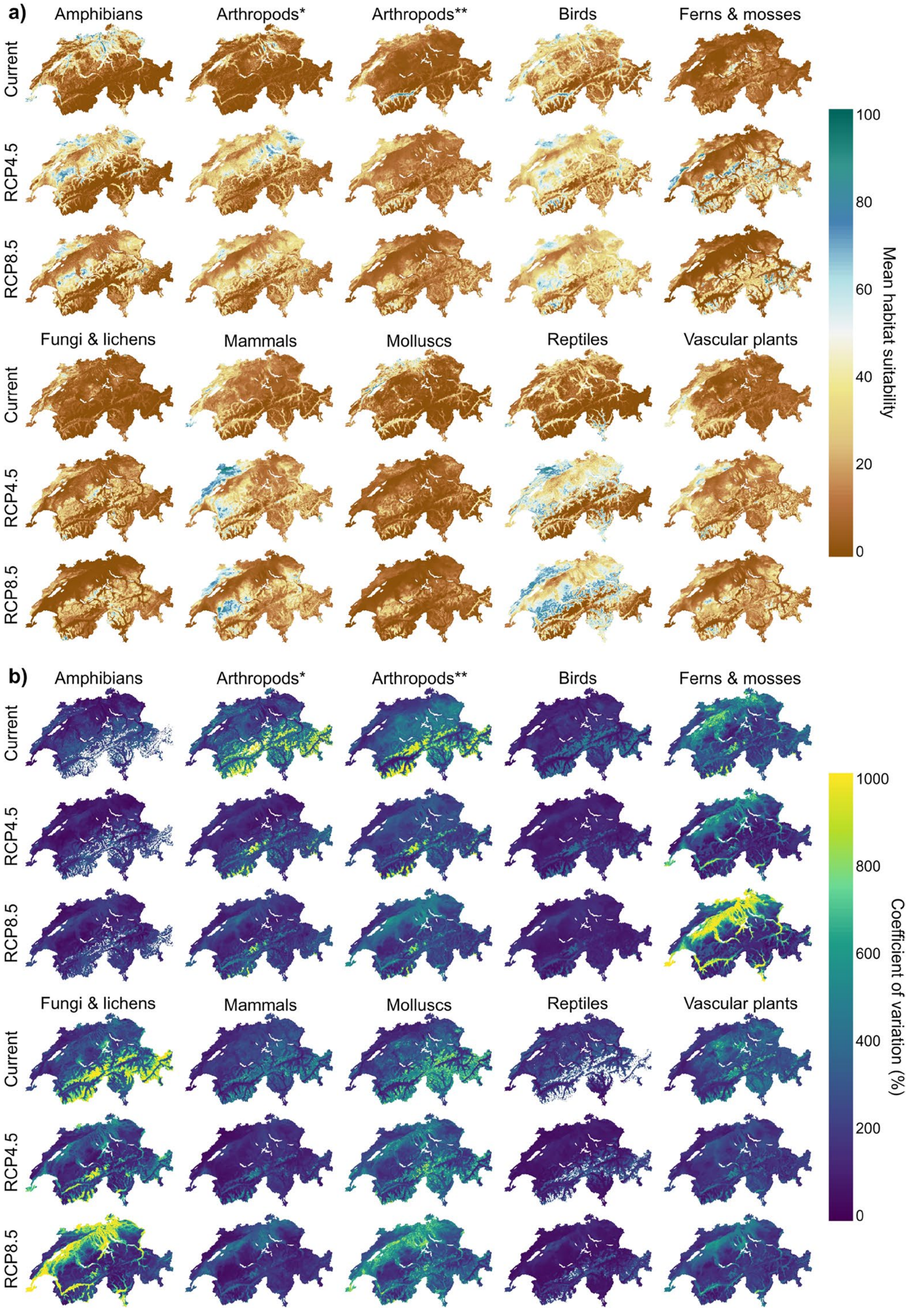

Fig. 2 SDMapCH (v1.3) (**a**) mean habitat suitability and (**b**) coefficient of variation maps for major taxonomic groups for current and future (2070–2099) periods under the low (RCP4.5) and high (RCP8.5) greenhouse gas scenarios. For each taxonomic group, mean values were rescaled to a 0–100 range for data storage purposes and improved readability, and coefficient of variation values were capped at 1,000 for readability. Arthropods*: aquatic arthropods including Ephemeroptera, Plecoptera, Trichoptera, and Odonata (EPTO). Arthropods**: All other arthropods excluding EPTO. Number of modelled species per group: amphibians (18), arthropods (2,285, including 296 EPTO) breeding birds (179), ferns and mosses (591), fishes (37), fungi and lichens (1,713), mammals (83), molluscs (168), reptiles (15), and vascular plants (2,410). For the use and validation of aggregated habitat suitability maps, please note: (i) The habitat suitability maps were primarily developed for research purposes and must be validated for case-specific conservation purposes; (ii) The maps represent aggregations of taxonomic groups for which species were not validated individually; therefore, interpretation should be complemented with additional taxonomic or faunistic expertise; (iii) The results of these models were not systematically validated by InfoSpecies; (iv) The maps should not be used directly for practical applications or management decisions without consultation with experts in the relevant taxonomic groups.

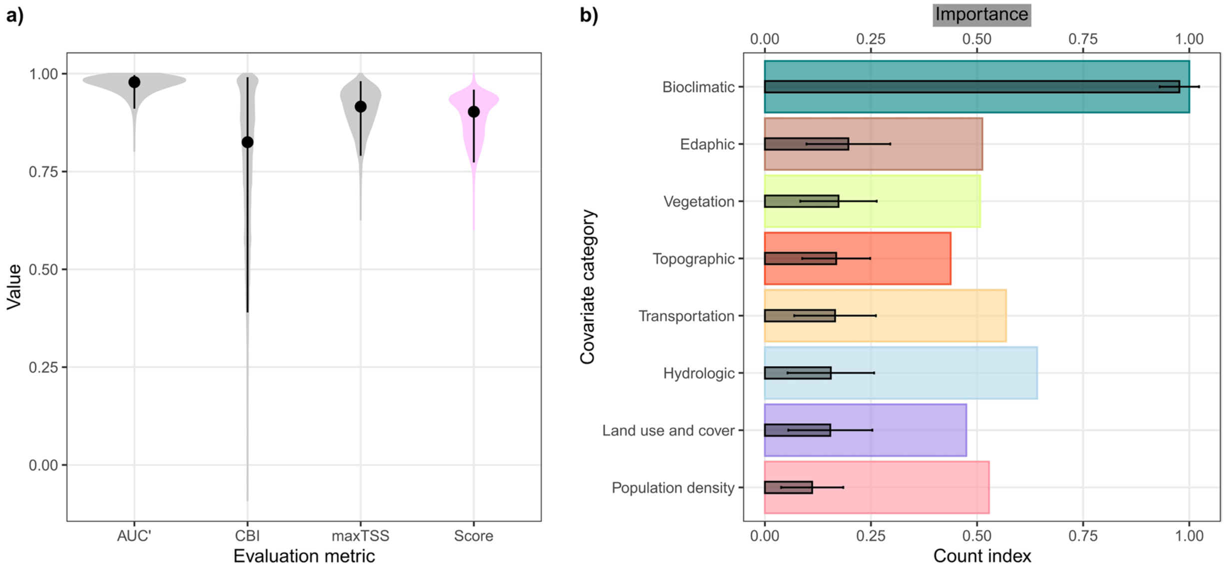

Fig. 3 SDMapCH (v1.3) model evaluation and covariate selection results. (**a**) Cross-validated model evaluation metrics (AUC': Somers' D Area Under the Curve, CBI: Continuous Boyce Index, maxTSS: maximized True Skill, and their average Score) for the 7,508 species and averaged across the five modelling algorithms (Generalized Linear Model, Generalized Additive Model, Maxnet, Random Forest, and Light Gradient Boosted Machine). Violin plots illustrate data distributions, showing the median (black dots) and the 2.5th to 97.5th percentile range (black lines). (**b**) Selection frequency and importance of the covariate categories for the 7,508 species. The covariate 'count index' (large colored bars) is the log of the count of covariates from a category selected in the final models divided by the overall number of covariates available in this category. The 'importance' (thin dark bars) is the average importance value of the covariates from a category in the final models across the 7,508 species and five algorithms.

Out of the 858 candidate habitat covariates available at the regional level, 645 were selected in at least one model. The global model output covariate was forced in all models and systematically outperformed the others in terms of relative importance (0.98 ± 0.05) (Fig. 3b). Among the other seven categories of regional-level covariates, 'hydrologic' covariates were the most often selected relative to the overall number of candidates in this category, whereas 'topographic' covariates were the least (Fig. 3b). Aside from the global model output, the 'edaphic' category had the highest average importance (0.20 ± 0.10), while the 'population density' category the lowest (0.11 ± 0.07) (Fig. 3b).

Averaged at the taxonomic group level, covariate category importance values varied among groups, except for bioclimatic covariates, which consistently exceeded 0.96 (on a scale from 0 to 1). Then, hydrologic covariates had the highest importance for arthropods (EPTO) (0.26), reptiles (0.20), and molluscs (0.19). Edaphic covariates reached their maximum importance values (0.19 to 0.21) for fungi & lichens, ferns & mosses, and molluscs, followed by vascular plants. Other covariate categories, including vegetation, topographic, transportation, and land use/cover, showed intermediate importance across all groups, with peaks of about 0.18 to 0.20 depending on the group. Population density covariates had consistently low importance across all groups, ranging from 0.05 for fishes to 0.14 for ferns & mosses. These patterns likely reflect differences in ecological requirements and habitat preferences among the major taxonomic groups assessed in this study. All detailed covariate selection and importance results are provided in the metadata.zip file that accompanies the maps on Dryad.

Integrity of output layers. Errors during the layer writing process can lead to unwanted artifacts in the output files. The integrity of SDMapCH output layers was checked by using a simple convolutional neural network (CNN) made of two convolutional blocks and two fully connected layers. The CNN was trained on a data set of about 5,000 image tiles (512 × 512 pixels), half of them with artifacts, and tested with an independent set of about 1,400 tiles per category. The classification accuracy on the test set was of 97%. All SDMapCH layers were checked by applying the CNN on 17 tiles sampled in a diagonal line across the layers. Layers with corrupted tiles were flagged, inspected manually, and re-written if necessary.

Data availability

The SDMapCH database (v1.3)[55] and associated metadata are openly available on Dryad (https://doi.org/10.5061/dryad.stqjq2cdk).

Code availability

For each version of SDMapCH, the specific setting file used for running N-SDM can be found in the metadata.zip file accompanying the maps on Dryad. N-SDM is available from Zenodo at https://zenodo.org/records/10604559, and GitHub at https://github.com/antadde/N-SDM, where complementary instructions for installation and example data are provided. The CNN model and code used for checking layers integrity are available from the GitHub Repository https://github.com/ManuelPopp/QuickCheckRasterWriteError.

Received: 11 June 2025; Accepted: 24 September 2025;

References

1. Franklin, J. *Mapping species distributions: spatial inference and prediction*. (Cambridge University Press, 2010).
2. Guisan, A., Thuiller, W. & Zimmermann, N. E. *Habitat suitability and distribution models, with applications in R*. (Cambridge University Press, 2017).
3. Peterson, A. T. *et al*. *Ecological niches and geographic distributions*. (Princeton University Press, 2011).
4. Ferrier, S. *et al*. *IPBES: The methodological assessment report on scenarios and models of biodiversity and ecosystem services*. (Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, 2016).
5. Guisan, A. *et al*. Predicting species distributions for conservation decisions. *Ecology Letters* **16**, 1424–1435 (2013).
6. Araujo, M. B. *et al*. Standards for distribution models in biodiversity assessments. *Science Advances* **5**, eaat4858 (2019).
7. Kays, R., McShea, W. J. & Wikelski, M. Born-digital biodiversity data: Millions and billions. *Diversity and Distributions* **26**, 644–648 (2020).
8. Jetz, W. *et al*. Essential biodiversity variables for mapping and monitoring species populations. *Nature Ecology & Evolution* **3**, 539–551 (2019).
9. Graham, E. M. *et al*. Climate change and biodiversity in Australia: a systematic modelling approach to nationwide species distributions. *Australasian Journal of Environmental Management* **26**, 112–123 (2019).
10. Keller, V. *et al*. *European breeding bird atlas 2: distribution, abundance and change*. (European Bird Census Council & Lynx Edicions, Barcelona, 2020).
11. Adde, A. *et al*. N-SDM: a high-performance computing pipeline for Nested Species Distribution Modelling. *Ecography*, e06540 (2023).
12. Guisan, A. *et al*. Spatially-nested species distribution models (N-SDM): an effective tool to overcome niche truncation for better inference and projections. *Journal of Ecology* **113**, 1588–1605 (2025).
13. Petitpierre, B. *et al*. Will climate change increase the risk of plant invasions into mountains? *Ecological Applications* **26**, 530–544 (2016).
14. Adde, A. *et al*. Projecting untruncated climate change effects on species' climate suitability: Insights from an alpine country. *Global Change Biology* **30**, e17557 (2024).
15. Chauvier-Mendes, Y. *et al*. Transnational conservation to anticipate future plant shifts in Europe. *Nature Ecology & Evolution* **8**, 1–13 (2024).
16. Vignali, S., Lörcher, F., Hegglin, D., Arlettaz, R. & Braunisch, V. Modelling the habitat selection of the bearded vulture to predict areas of potential conflict with wind energy development in the Swiss Alps. *Global Ecology and Conservation* **25**, e01405 (2021).
17. Brunner, C. *et al*. *Concurrent challenges for Switzerland: interfaces in climate action, biodiversity conservation, energy security and landscape preservation*. (ETH Zurich, 2024).
18. Egli, T., Bolliger, J. & Kienast, F. Evaluating ecosystem service trade-offs with wind electricity production in Switzerland. *Renewable and Sustainable Energy Reviews* **67**, 863–875 (2017).
19. Külling, N. *et al*. Nature's contributions to people and biodiversity mapping in switzerland: Spatial patterns and environmental drivers. *Ecological Indicators* **163**, 112079 (2024).
20. Ramel, C. *et al*. Integrating ecosystem services within spatial biodiversity conservation prioritization in the Alps. *Ecosystem Services* **45**, 101186 (2020).
21. Rey, P. *et al*. Predicting current and future spatial patterns of nature's contributions to people from species distribution models. *Ecological Indicators* **175**, 113528 (2025).
22. Zurell, D. *et al*. A standard protocol for reporting species distribution models. *Ecography* **43**, 1–17 (2020).
23. Hortal, J. *et al*. Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **46**, 523–549 (2015).
24. Meyer, C., Kreft, H. & Weigelt, P. Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecology Letters* **19**, 992–1006 (2015).
25. Beck, J., Ballesteros-Mejia, L., Nagel, P. & Kitching, I. J. Online solutions and the 'W allacean shortfall': what does GBIF contribute to our knowledge of species' ranges? *Diversity and Distributions* **19**, 1043–1050 (2013).
26. Moudry, V. & Simova, P. Influence of positional accuracy, sample size and scale on modelling species distributions: a review. *International Journal of Geographical Information Science* **26**, 2083–2095 (2012).
27. Braunisch, V. & Suchant, R. Predicting species distributions based on incomplete survey data: the trade-off between precision and scale. *Ecography* **33**, 826–840 (2010).
28. Qiao, H. *et al*. An evaluation of transferability of ecological niche models. *Ecography* **42**, 521–534 (2019).
29. Dépraz, A. *et al*. Swiss species occurrence data for controlled access by SPEED2ZERO (v1) (Swiss National Biodiversity Data and Information Centres – infospecies.ch. *GBIF* https://doi.org/10.15468/x9z635 (2025).
30. Klaus, G., Cordillot, F. & Künzle, I. *Gefährdete Arten und Lebensräume in der Schweiz: Synthese Rote Listen*. (BAFU, Bern, 2023).
31. Delarze, R., Gonseth, Y., Eggenberg, S. & Vust, M. *Guide des milieux naturels de Suisse - Ecologie, menaces, espèces caractéristiques*. (Rossolis, Bussigny, 2015).
32. Zizka, A. *et al*. CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* **10**, 744–751 (2019).
33. Boria, R. A., Olson, L. E., Goodman, S. M. & Anderson, R. P. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling* **275**, 73–77 (2014).
34. Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B. & Anderson, R. P. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38**, 541–545 (2015).
35. Steen, V. A., Tingley, M. W., Paton, P. W. C. & Elphick, C. S. Spatial thinning and class balancing: Key choices lead to variation in the performance of species distribution models with citizen science data. *Methods in Ecology and Evolution* **12**, 216–226 (2020).
36. Fourcade, Y., Engler, J. O., Rodder, D. & Secondi, J. Mapping species distributions with maxent using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *Plos One* **9**, 13 (2014).
37. Lamboley, Q. & Fourcade, Y. No optimal spatial filtering distance for mitigating sampling bias in ecological niche models. *Journal of Biogeography* **51**, 1783–1794 (2024).

38. Wisz, M. S. & Guisan, A. Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology* **9**, 1–13 (2009).
39. Barbet-Massin, M., Jiguet, F., Albert, C. H. & Thuiller, W. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* **3**, 327–338 (2012).
40. Steen, B., Broennimann, O., Maiorano, L. & Guisan, A. How sensitive are species distribution models to different background point selection strategies? A test with species at various equilibrium levels. *Ecological Modelling* **493**, 110754 (2024).
41. Külling, N. *et al.* SWECO25: a cross-thematic raster database for ecological research in Switzerland. *Scientific Data* **11**, 1–21 (2024).
42. Karger, D. N. *et al.* Climatologies at high resolution for the earth's land surface areas. *Scientific Data* **4**, 1–20 (2017).
43. van Vuuren, D. P. *et al.* The representative concentration pathways: an overview. *Climatic Change* **109**, 5–31 (2011).
44. Adde, A. *et al.* Too many candidates: Embedded covariate selection procedure for species distribution modelling with the covsel R package. *Ecological Informatics* **75**, 102080 (2023).
45. Warton, D. I., Renner, I. W. & Ramp, D. Model-based control of observer bias for the analysis of presence-only data in ecology. *Plos One* **8**, 9 (2013).
46. Chauvier, Y. *et al.* Novel methods to correct for observer and sampling bias in presence-only species distribution models. *Global Ecology and Biogeography* **30**, 2312–2325 (2021).
47. Bonnet-Lebrun, A. S., Karamanlidis, A. A., Hernando, M. D., Renner, I. & Gimenez, O. Identifying priority conservation areas for a recovering brown bear population in Greece using citizen science data. *Animal Conservation* **23**, 83–93 (2020).
48. McCullagh, P. & Nelder, J. A. *Generalized linear models. 2nd edition*. (Chapman and Hall, London, 1989).
49. Hastie, T. J. in *Statistical models in S*. 249-307 (Routledge, London, 2017).
50. Phillips, S. J., Anderson, R. P., Dudik, M., Schapire, R. E. & Blair, M. E. Opening the black box: an open-source release of Maxent. *Ecography* **40**, 887–893 (2017).
51. Breiman, L. Random forests. *Machine Learning* **45**, 5–32 (2001).
52. Ke, G. L. *et al.* LightGBM: A highly efficient gradient boosting decision tree. *Adv Neur In* **30** (2017).
53. Somers, R. H. A new asymmetric measure of association for ordinal variables. *American sociological review*, 799-811 (1962).
54. Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling* **199**, 142–152 (2006).
55. Adde, A. *et al.* SDMapCH (v1.3): a Comprehensive database of modelled species habitat suitability maps for Switzerland. *Dryad* https://doi.org/10.5061/dryad.stqjq2cdk (2025).
56. Guillera-Arroita, G. *et al.* Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography* **24**, 276–292 (2015).

Acknowledgements

This study is part of the SPEED2ZERO Joint Initiative (https://speed2zero.ethz.ch/), which received support from the ETH-Board under the Joint Initiatives scheme. We gratefully acknowledge financial support from the Federal Office for the Environment (FOEN) through the Action Plan of the Swiss Biodiversity Strategy for financing the ValPar.CH (https://www.valpar.ch/) and SwissCatchment projects, within the context of which the N-SDM pipeline and preliminary SDMapCH versions 1.1 and 1.2 were developed. The Swiss Species Information Center InfoSpecies (www.infospecies.ch) supplied Swiss-level species occurrence data and expertise on species' ecology, and we acknowledge their support regarding the database. This research was enabled in part by the support provided by the Scientific Computing and Research Unit of Lausanne University (www.unil.ch/ci/dcsr). We thank two anonymous reviewers for their comments on the manuscript. Finally, this work would not have been possible without the dedication of many individuals who recorded and reported species occurrences. We sincerely thank all contributors, ranging from citizen scientists to taxonomic specialists, for their valuable efforts and expertise, which were essential to the development of this database.

Author contributions

Antoine Adde co-led the conceptualization of the study, led the development of the database and its validation, and wrote the original draft. Pierre-Louis Rey contributed to the conceptualization of the study, processing of species and covariate data, and review and editing of the manuscript. Nathan Külling contributed to the conceptualization of the study, processing of covariate data, figure preparation, and review and editing of the manuscript. Yohann Chauvier-Mendes contributed to the processing of species data and review and editing of the manuscript. Fabian Fopp contributed to the conceptualization of the study, processing of covariate data, and review and editing of the manuscript. Manuel R. Popp contributed to database validation and review and editing of the manuscript. Olivier Broennimann contributed to the conceptualization of the study, processing of covariate data, and review and editing of the manuscript. Blaise Petitpierre contributed to the conceptualization of the study, database validation, and review and editing of the manuscript. Nicolas Strebel contributed to database validation and review and editing of the manuscript. Andrin Gross contributed to database validation and review and editing of the manuscript. Silvia Stofer contributed to species data access, database validation, and review and editing of the manuscript. Anthony Lehmann contributed to the conceptualization of the study and review and editing of the manuscript. Niklaus E. Zimmermann contributed to the conceptualization of the study and review and editing of the manuscript. Loïc Pellissier contributed to the conceptualization of the study and review and editing of the manuscript. Antoine Guisan co-led the funding acquisition, supervision, and conceptualization of the study, and contributed to the review and editing of the manuscript. Florian Altermatt co-led the funding acquisition, supervision, and conceptualization of the study, and contributed to the review and editing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.A., A.G. or F.A.

Reprints and permissions information is available at www.nature.com/reprints.