https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

Open Access Earth System Science Data Discussions

A global long-term ecoregion-level dataset of biodiversity indicators for terrestrial vertebrates under climate and habitat change

Shijun Zheng[1, #], Hui Wu[2, #], Le Yu[1,3,*], Keping Ma[2]

[1] Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing 100084, China

[2]Northeast Asia Biodiversity Research Center, College of Forestry, Northeast Forestry University, Harbin 150040, China

[3] Ministry of Education Ecological Field Station for East Asian Migratory Birds, Beijing 100084, China.

#Contributed equally

[*] Corresponding author: leyu@tsinghua.edu.cn

Keywords: biodiversity; terrestrial vertebrates; AOH; ecoregion

Abstract

Climate and habitat change are reshaping global biodiversity, yet long-term, spatially explicit datasets that capture these dynamics across ecologically meaningful regions remain limited. Here, we present a global spatiotemporal dataset of ecoregion-level biodiversity indicators for terrestrial vertebrates derived from species-specific Area of Habitat (AOH). By integrating species occurrence records, expert-derived range maps, climate data, and temporally explicit habitat maps, we reconstructed distributions for 19146 species of amphibians, birds, mammals, and reptiles. We then derived four complementary biodiversity indicators—species richness, threatened species richness, species endemism, and AOH density—for global terrestrial ecoregions at five time points (1990, 2020, 2030, 2050, and 2100), with future projections under SSP245 and SSP585. Validation at both species and dataset levels demonstrated good performance. Predicted AOH consistently assigned higher

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

Open Access Earth System Science Data Discussions

suitability to independent occurrence records than to the broader range background, and estimated richness closely reproduced the major spatial patterns of occurrence-derived richness across ecoregions. In 2020, the four indicators showed distinct but broadly concordant spatial patterns, with the highest overall values concentrated in tropical ecoregions. Future projections revealed marked spatial heterogeneity, with declines in multiple indicators concentrated in many tropical ecoregions by 2100 and generally becoming stronger under SSP585. This dataset provides a globally consistent and ecologically meaningful resource for biodiversity monitoring, macroecological analysis, and conservation assessment under ongoing climate and habitat change. The dataset supporting this study is publicly available at https://doi.org/10.5281/zenodo.20119261.

1 Introduction

Biodiversity is undergoing rapid and widespread change under the combined pressures of climate change and land-use transformation, with profound implications for ecosystem functioning and human well-being (Díaz et al., 2019; Isbell et al., 2017). Despite rapid advances in biodiversity science, our ability to consistently track global biodiversity change across space and time remains limited (Pettorelli et al., 2014; Jetz et al., 2019). Multiple lines of evidence indicate that human activities are driving substantial declines in species diversity and altering community composition across terrestrial ecosystems (Ipbes, 2019; Díaz et al., 2019; Newbold et al., 2015). Land-use change and climate change have been identified as dominant drivers of biodiversity loss globally, reshaping species distributions and ecosystem structure (Pereira et al., 2010; Tilman et al., 2017). These changes not only threaten species persistence but also impair key ecosystem functions, including productivity, carbon storage, and climate regulation (Cardinale et al., 2012; Hooper et al., 2012). In this context, robust, spatially explicit, and temporally consistent biodiversity data are essential for understanding global change impacts and supporting conservation planning.

Despite increasing data availability, current biodiversity assessments remain

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

limited by fragmentation of data sources and methodological inconsistencies. Species occurrence records, such as those from the Global Biodiversity Information Facility (GBIF), provide extensive observational data but are affected by strong spatial sampling biases (Beck et al., 2014; Meyer et al., 2016), whereas expert-derived range maps, such as those from the International Union for Conservation of Nature (IUCN), offer global coverage but often lack fine-scale ecological realism and may overestimate species' actual distributions (Maiorano, 2011). Species distribution models (SDMs) have substantially improved the prediction of species distributions by integrating environmental variables with occurrence data, yet they are typically applied to individual species or limited regions rather than generating harmonized global datasets (Elith and Leathwick, 2009; Franklin, 1945). As a result, these data sources are rarely integrated within a consistent framework, leading to fragmented and often biased representations of species distributions and limiting the ability to capture the joint effects of climate suitability and habitat availability.

Another major limitation is the lack of harmonized, long-term biodiversity indicators at ecologically meaningful spatial units. Existing global datasets are often structured at grid-cell or administrative scales, which may not align with ecological processes or species assemblage patterns. Ecoregions, defined as areas of relatively homogeneous environmental conditions and species composition, provide a more ecologically relevant unit for summarizing biodiversity patterns (Olson et al., 2001). However, globally consistent datasets that quantify multiple dimensions of biodiversity across ecoregions and through time remain scarce, particularly those integrating multiple taxa and future environmental scenarios. The concept of Area of Habitat (AOH) provides a promising framework to address these challenges. AOH refines species distributions by combining habitat preferences with environmental suitability, thereby identifying areas that are both ecologically suitable and climatically viable (Brooks et al., 2019; Lumbierres et al., 2022). Compared with traditional range maps, AOH

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

reduces commission errors and improves the ecological realism of species distribution estimates (Dahal et al., 2021). However, most existing AOH-based studies have focused on present-day patterns, specific regions, or limited taxonomic groups, and have not been extended to produce globally harmonized, long-term biodiversity indicators.

Here, we present a global, spatiotemporal dataset of ecoregion-level biodiversity indicators for terrestrial vertebrates derived from species-specific AOH. We integrate species occurrence records, expert range constraints, climate projections from CMIP6, and globally consistent habitat maps to generate AOH distributions for 19146 species of amphibians, birds, mammals, and reptiles at five time points (1990, 2020, 2030, 2050, and 2100), with future projections provided under SSP245 and SSP585. Based on these distributions, we derive four complementary biodiversity indicators—species richness, threatened species richness, species endemism, and AOH density—aggregated at the ecoregion level. To our knowledge, this represents one of the first globally consistent, long-term biodiversity indicator datasets at the ecoregion scale. This dataset provides a harmonized and process-informed representation of global biodiversity patterns and their temporal dynamics, enabling consistent comparisons across regions, taxa, and time, and offering a robust foundation for biodiversity monitoring, macroecological research, and conservation decision-making under ongoing global change.

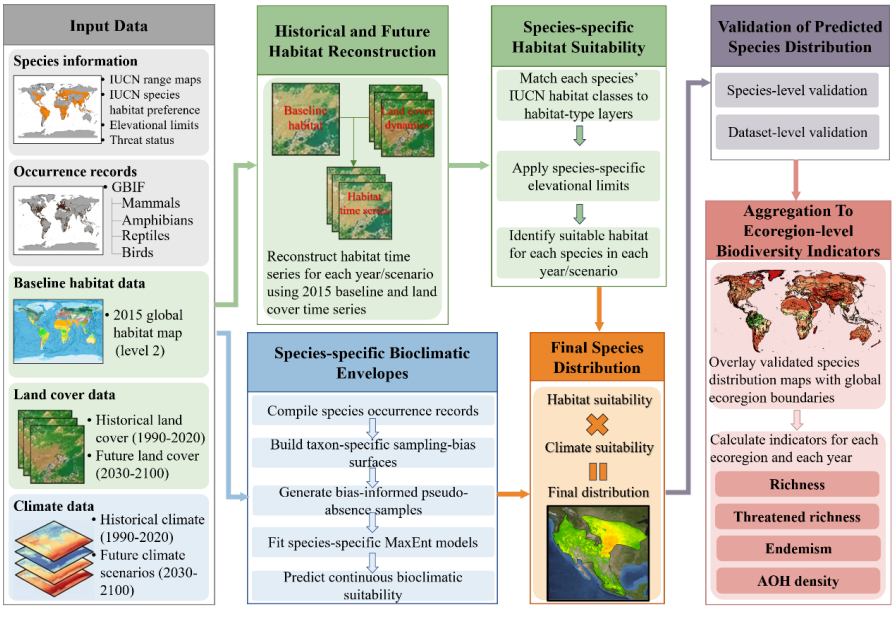
2 Material and methods

2.1 Overall workflow

We developed an integrated workflow to generate a global long-term dataset of ecoregion-level biodiversity indicators for terrestrial vertebrates under climate and habitat change. The workflow combined species range and occurrence data, species-specific ecological attributes, climate data, terrestrial habitat maps, and a global ecoregion framework to reconstruct species distributions through time and to summarize biodiversity change at the ecoregion level (Fig. 1).

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

First, we compiled species range maps, habitat associations, elevational limits, conservation status, and occurrence records for amphibians, birds, mammals, and reptiles. We also assembled climate data, a baseline terrestrial habitat map, temporally reconstructed habitat layers, and the ecoregion data as the spatial basis for aggregation. These datasets provided the biological and environmental inputs for subsequent modelling. Second, we estimated species-specific climatic suitability using occurrence-based MaxEnt models and derived habitat suitability by linking IUCN habitat associations and elevational limits to temporally explicit habitat maps. Climatic suitability and habitat suitability were then integrated within expert-defined geographic ranges to produce species-specific Area of Habitat (AOH) maps for historical and future periods. Finally, species-level AOH maps were overlaid with ecoregion boundaries to calculate four biodiversity indicators for each ecoregion, year, and scenario: species richness, threatened species richness, species endemism, and AOH density. Together, these steps generated a harmonized spatiotemporal dataset for assessing biodiversity patterns and long-term changes across global terrestrial ecoregions.



https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

Fig. 1 Schematic workflow for constructing the global ecoregion-level biodiversity indicator dataset based on species-specific habitat and climatic suitability.

2.2 Data acquisition

2.2.1 Species range and occurrence data

By combining expert-derived range constraints, habitat associations, and occurrence records, we established a unified species database to support subsequent modelling of habitat suitability, climatic suitability, and final species distributions. We first compiled taxonomic information and species-specific ecological attributes from the International Union for Conservation of Nature (IUCN) Red List using the R package rredlist (Gearty and Chamberlain, 2022). For each species, we retrieved the accepted scientific name, habitat preference information, elevational limits and conservation status. Habitat associations provided by IUCN were recorded as suitable, marginally suitable, or unsuitable and were later used to translate coarse expert range boundaries into habitat-filtered distributions. Elevational limits were used as an additional ecological constraint to exclude environmentally unrealistic portions of otherwise suitable habitat. Together, these species-level attributes formed the basis for refining broad geographic ranges into more ecologically realistic AOH estimates.

Species range polygons were compiled from the most widely used global expert range datasets for terrestrial vertebrates. Range maps for amphibians, mammals, and reptiles were obtained from the IUCN Red List spatial database (Brooks et al., 2019), whereas bird range maps were obtained from BirdLife International (Birdlife, 2015). This combination follows common practice in global biodiversity and AOH assessments, because BirdLife provides the most comprehensive and standardized range data for birds, while IUCN is the principal source for the other terrestrial vertebrate groups. These polygon datasets were used as the primary geographic constraints within which habitat and climate suitability were later evaluated. To ensure consistency among species range maps, habitat layers, and ecological metadata, we

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

applied a series of screening and filtering procedures before modelling. First, only polygons classified as "Extant" or "Probably Extant" were retained, so that the analysis represented current known or likely occupied distributions rather than historical, uncertain, or introduced occurrences. Second, species lacking key ecological information required for AOH derivation were excluded. Such missing information included absent or incomplete habitat associations, unavailable elevational limits where needed for habitat filtering, or unresolved taxonomic inconsistencies that prevented linkage among range, habitat, and occurrence datasets. Third, we excluded species whose ecology could not be adequately represented using terrestrial land-cover-based habitat masks. This mainly concerned taxa strongly associated with caves, subterranean environments, or other habitat systems that are not reliably characterized in global land-cover products. Fourth, species were removed when inconsistencies among their range polygons, habitat preferences, and habitat maps resulted in no remaining suitable habitat after filtering, because AOH could not be meaningfully derived in such cases. Because the objective of this dataset was to characterize terrestrial biodiversity, we further excluded species associated with non-terrestrial systems based on the IUCN "systems" classification. Specifically, species categorized as "marine", "terrestrial + marine", "freshwater", or "freshwater + marine" were not included in the final species pool. This step ensured that the dataset remained focused on taxa whose distributions can be reasonably interpreted using terrestrial habitat and climate variables.

In addition to expert-derived range and trait information, we assembled species occurrence records to support climatic niche modelling and subsequent validation. Occurrence data were downloaded from the Global Biodiversity Information Facility (GBIF) (Secretariat, 2020), with species names matched to the accepted scientific names compiled from IUCN. We retrieved all georeferenced occurrence records available between 1990 and 2025 for each species. This temporal window was selected to maximize data availability while remaining broadly consistent with the period

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

represented by contemporary climate and land-cover information used in the analysis. Because birds often exhibit seasonally distinct distributions, we distinguished between breeding-associated and non-breeding-associated occurrence records when possible. For birds, we therefore constructed two occurrence datasets: a breeding-season dataset, including records located within or near breeding and resident range polygons, and a non-breeding dataset, including records located within or near non-breeding and resident range polygons. This distinction allowed seasonal climatic suitability to be represented more realistically for migratory and partially migratory bird species. For amphibians, mammals, and reptiles, such separation was not performed because temporally explicit breeding and non-breeding range maps are generally unavailable at the global scale for these groups (Smith et al., 2025). Raw GBIF records were subjected to multiple quality-control steps prior to use. Fossil records were removed because they do not represent contemporary species distributions. We also excluded records associated with problematic coordinate flags, including presumed negated longitude, presumed negated latitude, suspicious coordinate reprojection, invalid coordinate precision, rounded coordinates, and invalid geodetic datum. These filters were applied to reduce the influence of geolocation errors and spatial uncertainty that could distort model calibration. To further minimize the inclusion of erroneous, exotic, captive, or otherwise implausible records, we retained only records falling within a 500-km buffer around each species' expert-derived range polygon. This buffer-based screening was used to remove extreme outliers while preserving records located near mapped range boundaries, which may still reflect valid occurrences despite uncertainty in expert range delineation. Such filtering is particularly important in global occurrence datasets, where some records may correspond to zoo specimens, misidentifications, translocated populations, or other occurrences outside the expected native distribution (Zizka et al., 2019).

After taxonomic harmonization and quality filtering, the resulting species database

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

contained, for each retained species, a linked set of expert range constraints, habitat associations, elevational limits, conservation status, system classification, and cleaned occurrence records. These data constituted the biological foundation of the modelling framework and were subsequently integrated with land-cover and climate datasets to estimate species-specific AOH and derive ecoregion-level biodiversity indicators.

2.2.2 Current and future climate data

Climate data were derived from the CMIP6-based products provided through the IPCC AR6 Interactive Atlas (Change, 2013). We initially considered five climatic variables that are widely recognized as important in defining species' bioclimatic envelopes: mean annual temperature, annual maximum temperature, annual minimum temperature, mean annual precipitation, and consecutive dry days. Because the three temperature variables were strongly correlated, we reduced redundancy by retaining mean annual temperature and deriving annual temperature range as the difference between maximum and minimum temperature (Smith et al., 2025). These two temperature-related variables, together with mean annual precipitation and consecutive dry days, were used as climatic predictors in subsequent modelling. The baseline climate layers covered the period 1990–2014, while future climate layers covered 2015–2100 under two CMIP6 scenarios, SSP2-4.5 (SSP245) and SSP5-8.5 (SSP585). These two scenarios were selected to represent contrasting future climate trajectories, with SSP245 reflecting an intermediate pathway and SSP585 representing a high-emissions pathway. Using both scenarios allowed us to assess biodiversity responses under moderate and more severe future climate change.

Because the occurrence records used in this study spanned 1990–2025, the historical climate layers alone did not fully match the temporal range of the biological observations used for model calibration. To reduce this temporal mismatch, we extended the baseline climate period by incorporating SSP245 layers for 2015–2025 and used the combined 1990–2025 period to represent contemporary climatic

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

conditions. This approach is justified because differences among CMIP6 scenarios are relatively small in the near term, while SSP2-4.5 represents an intermediate emissions pathway and therefore provides a pragmatic extension of recent climate conditions for the period immediately following the historical baseline (Masson-Delmotte et al., 2021). Using this extended baseline improved temporal consistency between occurrence records and environmental predictors.

2.2.3 Terrestrial habitat map and temporal reconstruction

To represent habitat suitability in a manner consistent with species-specific IUCN habitat associations, we used the global terrestrial habitat map developed by Jung et al. (2020). This product provides a spatially explicit global classification of terrestrial habitat types for the year 2015 at 100 m resolution and was generated by refining the Copernicus 2015 land-cover dataset through a global decision-tree framework. The map was developed and evaluated using occurrence information from 828 vertebrate species, including 35152 point records, 8181 polygon records, and 6026 sampling sites. Reported average balanced accuracy was 0.77 ± 0.14 for Level 1 habitat classes and 0.71 ± 0.15 for Level 2 classes, indicating good performance for broad-scale habitat characterization. An important advantage of this dataset is that its habitat classes are directly aligned with the IUCN habitat classification scheme, allowing habitat preferences recorded for each species to be linked explicitly to mapped habitat conditions. We used the Level 2 habitat classes because they provide finer ecological discrimination than Level 1 while remaining globally consistent. To match the spatial resolution of the broader modelling framework, the original 100 m habitat map was aggregated to a 10 km × 10 km grid, in which each cell stores the proportional coverage of each habitat type. In total, 58 Level 2 terrestrial habitat classes were included in the analysis (Table S1).

Because this habitat map represents only a single reference year (2015), it could not be used directly to support the production of long-term biodiversity indicators

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

257 spanning both historical and future periods. We therefore extended this static baseline
258 into a temporally continuous habitat dataset by combining it with time-varying land-
259 cover information for the past and projected vegetation information for the future
260 (Smith et al., 2025). Historical habitat dynamics were reconstructed using the From-
261 GLC Plus land-cover dataset (Yu et al., 2025), while future habitat changes were
262 inferred from plant functional type (PFT) projections (Chen et al., 2022). Following the
263 general approach of previous studies, these datasets were used to translate annual land-
264 cover and vegetation changes into corresponding habitat-type changes relative to the
265 2015 baseline map. This procedure allowed us to generate temporally consistent habitat
266 maps for the full study period while preserving compatibility with the IUCN habitat
267 classification system. The detailed rules are provided in Text S1.

268 ***2.2.4 Terrestrial ecoregion framework***

269 To aggregate biodiversity patterns at an ecologically meaningful spatial unit, we
270 used the WWF Terrestrial Ecoregions of the World (TEOW) framework (Olson et al.,
271 2001). This framework originally includes 825 terrestrial ecoregions, together with one
272 ice and one lake unit. Although the latter two are not part of the standard terrestrial
273 biome system, they were retained in this study because some terrestrial vertebrate
274 species in our dataset had suitable habitat or distributional overlap within these regions.
275 Compared with regular grid cells or administrative units, ecoregions provide a more
276 ecologically relevant basis for summarizing biodiversity patterns, as they better reflect
277 broad biogeographic structure, environmental heterogeneity, and species turnover.

278 For biome-level analyses and visualization, original 14 biomes were further
279 reclassified into eight broader biome groups (Fig. 2; Table S3) to simplify interpretation
280 while retaining the major ecological contrasts among global terrestrial systems. The
281 eight biome groups used in this study were: tropical forests, temperate forests, boreal
282 forests, tropical grasslands, temperate grasslands, tundra, mediterranean and deserts &
283 xeric shrublands. The number of ecoregions within each reclassified biome is shown in

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

Fig. 2. This ecoregion framework was used consistently throughout the study for spatial aggregation, biome-level summarization, and interregional comparison of biodiversity indicators.

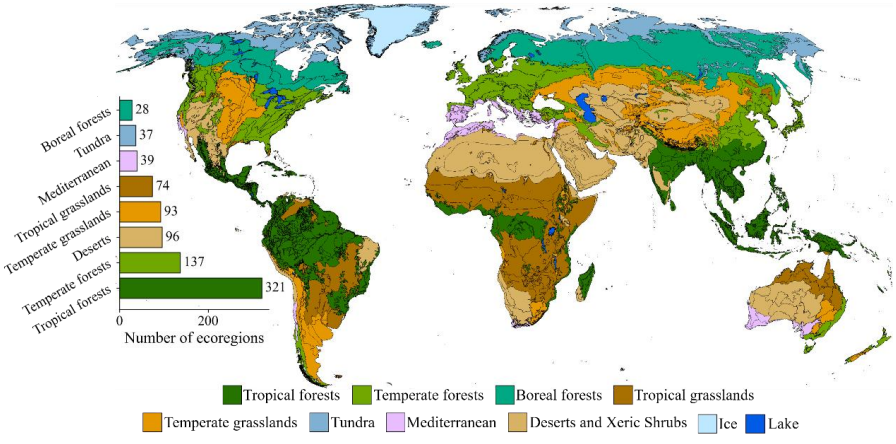


Fig. 2 Global terrestrial ecoregions and their reclassification into nine biomes used in this study. *The original terrestrial biomes were reclassified into eight broader biome types for analysis (Table S3). The inset bar chart indicates the number of ecoregions within each biome type.*

2.3 AOH modelling framework

2.3.1 Climatic suitability modelling

Species-specific climatic suitability was estimated using occurrence-based species distribution models. For each species, cleaned occurrence records were combined with gridded climate predictors to characterize its bioclimatic envelope, that is, the range of climatic conditions associated with known observations. We used MaxEnt to model climatic suitability because it is well suited for presence-background data and has been widely applied in large-scale biodiversity modelling (Phillips and Dudík, 2008).

To ensure sufficient data support for model calibration, climatic suitability was modelled only for species with at least 10 effective occurrence records after data

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

302 cleaning and spatial filtering. Here, effective occurrences were defined as unique
303 records located within a 500 km buffer around the expert-derived species range after
304 removing duplicates with identical spatial grids. This threshold was adopted to reduce
305 model instability for species represented by very few records. The distribution of
306 effective sample sizes varied substantially among taxa (Fig. S1), with birds generally
307 having larger sample sizes than amphibians, mammals, and reptiles. In total, 19146
308 terrestrial vertebrate species were included—comprising 3324 amphibians, 8119 birds,
309 2844 mammals, and 4859 reptiles. For species meeting the sample-size criterion,
310 climatic suitability was modelled using four predictors: mean annual temperature,
311 annual temperature range, mean annual precipitation, and consecutive dry days. These
312 variables were selected to represent major dimensions of thermal and moisture
313 conditions while limiting collinearity among predictors (Smith et al., 2025). All
314 modelling was conducted on the common 10 km grid used throughout the study.

315 MaxEnt models were fitted using species occurrence records together with taxon-
316 specific background points that accounted for spatial sampling bias in the global
317 occurrence data. Because occurrence records are unevenly distributed in space,
318 background sampling was constrained to areas that were both geographically accessible
319 and environmentally relevant to each species, thereby reducing artefacts caused by
320 highly biased observation effort (Lobo and Tognelli, 2011). Details of the background-
321 generation procedure are provided in Text S2. The MaxEnt models were run using
322 default parameter settings. Model outputs were retained as continuous climatic
323 suitability surfaces ranging from 0 to 1, rather than being converted immediately to
324 binary presence–absence maps. These values were generated using the logistic output
325 function, which transformed raw MaxEnt predictions into relative suitability scores for
326 each grid cell. This choice preserved more information on the relative climatic
327 favorability of each grid cell and avoided the additional uncertainty introduced by
328 threshold selection at the modelling stage (Guillera-Arroita et al., 2015; Santini et al.,

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

2021).

2.3.2 Habitat suitability mapping

Habitat suitability was mapped for each species by linking its IUCN habitat associations to the temporally explicit habitat raster stacks described in Text S1. These raster stacks quantify, for each 10 km × 10 km grid cell, the proportional cover of 58 Level 2 habitat classes that are directly aligned with the IUCN habitat classification system. For species s in grid cell i, habitat suitability was calculated as:

$$H_{s,i} = \sum_{k=1}^{58} p_{i,k} \cdot w_{s,k} \quad (1)$$

where $p_{i,k}$ is the proportional cover of habitat class k in grid cell i, and $w_{s,k}$ is the habitat-affinity weight of species s for habitat class k. Habitat classes classified by IUCN as suitable were assigned a value of 1, whereas habitat classes classified as marginally suitable or unsuitable were assigned a value of 0. Under this formulation, $H_{s,i}$ ranges from 0 to 1 and represents the proportion of the grid cell covered by habitat considered suitable for species s. To further constrain habitat suitability ecologically, species-specific elevational limits were applied using the IUCN lower and upper elevation bounds. Grid cells falling outside the documented elevational range of a species were assigned a habitat suitability of zero.

This procedure was applied to each species and each historical or future habitat layer to generate temporally explicit habitat-suitability surfaces, which were subsequently combined with climatic suitability and expert-derived range constraints to derive species-specific AOH.

2.3.3 Construction of species-specific AOH

Species-specific AOH was derived by integrating climatic suitability and habitat suitability within each species' expert-defined geographic range. Climatic suitability was represented by the continuous MaxEnt output, with values ranging from 0 to 1, while habitat suitability was quantified as the proportion of each grid cell covered by habitat types classified as suitable for the species after applying elevational constraints.

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026


Open Access Earth System Science Data Discussions

356 To ensure that predictions did not extend beyond the known distributional limits of a
357 species, both surfaces were masked by the corresponding expert range polygon before
358 combination. For species s in grid cell i, AOH was calculated as:

359 $$AOH_{s,i} = R_{s,i} \times C_{s,i} \times H_{s,i} \quad (2)$$

360 where $R_{s,i}$ is the range mask of species in grid cell i ($R_{s,i}$ = 1 if the cell falls within
361 the species' expert range and 0 otherwise), $C_{s,i}$ is the climatic suitability score, and
362 $H_{s,i}$ is the habitat suitability value. Under this formulation, AOH values also range
363 from 0 to 1 and represent the joint suitability of climate and habitat within the species'
364 known range. This approach retains the continuous information contained in both
365 climatic and habitat suitability surfaces, rather than imposing an arbitrary threshold at
366 this stage. As a result, higher AOH values indicate grid cells where both climatic
367 conditions and habitat availability are more favorable for the species, whereas low
368 values indicate partial limitation by one or both components. Cells outside the expert
369 range or lacking either climatic or habitat suitability receive values of zero.

370 The above procedure was applied to each species for all historical and future
371 periods, producing temporally explicit AOH maps that form the basis of the long-term
372 species distribution dataset. Fig. 3 illustrates this workflow using *Reithrodontomys*
373 *megalotis* as an example, showing how occurrence records, climate predictors, habitat
374 suitability, and range constraints were integrated to generate a time series of species
375 distributions.

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

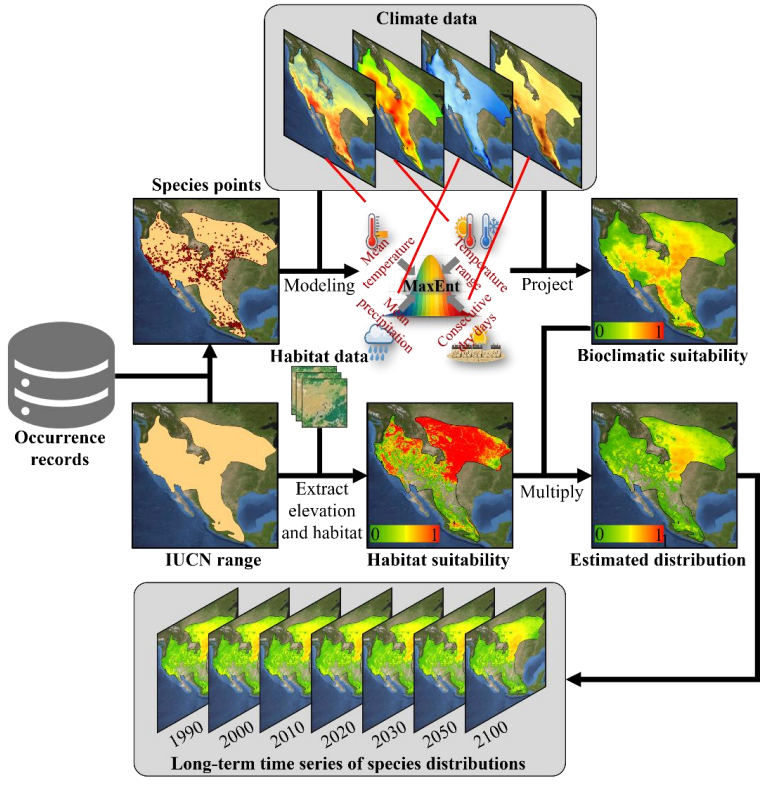


Fig. 3 Estimating species distribution using Reithrodontomys megalotis as an example.

2.4 Calculation of ecoregion-level biodiversity indicators

To characterize biodiversity patterns across ecoregions, we derived four complementary indicators from species-specific AOH maps: species richness, threatened species richness, species endemism, and AOH density (Fig. 4). For each year and scenario, these indicators were calculated independently for every ecoregion based on the spatial overlap between the ecoregion boundary and the AOH of each species.

Species richness was defined as the number of species with non-zero AOH within a given ecoregion. Thus, a species was considered present in ecoregion *e* when its AOH overlapped the ecoregion by any non-zero area. Under this definition, richness reflects

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

the total number of species for which climatically and habitat-suitable area occurs within the focal ecoregion, regardless of the magnitude of that overlap. Threatened species richness was calculated in the same manner, but restricted to species classified by the IUCN Red List as Vulnerable (VU), Endangered (EN), or Critically Endangered (CR). This indicator therefore represents the number of species of elevated conservation concern occurring within each ecoregion. Species endemism was designed to capture the degree to which species are spatially restricted to individual ecoregions. For each species occurring in a given ecoregion, we calculated the proportion of its total AOH located within that ecoregion, and then summed these proportions across all species present. Formally, endemism for ecoregion *e* was calculated as

$$Endemism_e = \sum_{i=1}^{S_e} \frac{AOH_{i,e}}{AOH_{i,global}} \quad (3)$$

where $AOH_{i,e}$ is the AOH area of species i within ecoregion e. $AOH_{i,global}$ is the total global AOH area of species i, and S_e is the number of species present in ecoregion e. Under this formulation, species entirely confined to one ecoregion contribute 1, whereas widely distributed species contribute less than 1 depending on the fraction of their AOH occurring within the focal ecoregion. AOH density was used to represent the intensity of habitat occupancy within each ecoregion. It was calculated as the sum of the AOH areas of all species within the ecoregion divided by the total area of that ecoregion:

$$AOH\ density_e = \frac{\sum_{i=1}^{S_e} AOH_{i,e}}{Area_e} \quad (4)$$

where $Area_e$ is the total area of ecoregion e. This metric integrates the cumulative amount of suitable habitat across species while standardizing for ecoregion size, allowing comparison among ecoregions with different spatial extents.

Together, these four indicators capture complementary dimensions of biodiversity: total species occurrence, concentration of threatened species, spatial uniqueness of

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

species distributions, and the relative abundance of suitable habitat within ecoregions.

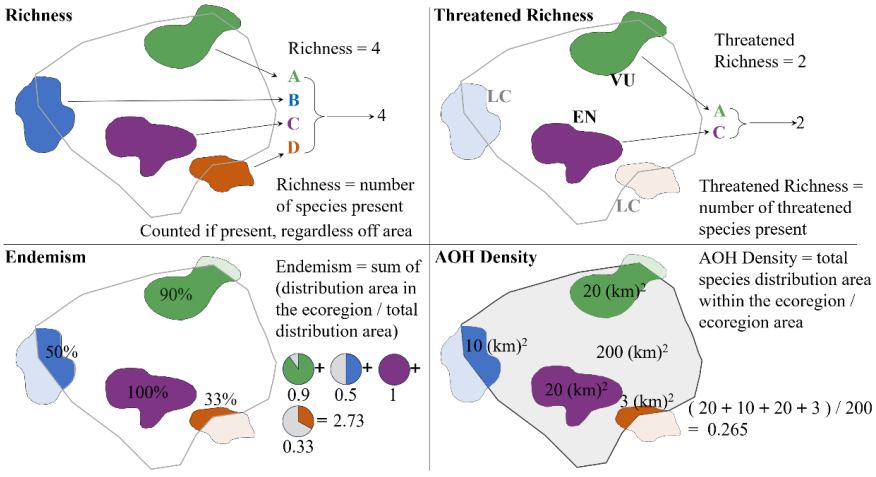


Fig. 4 Conceptual framework for calculating ecoregion-level richness, threatened richness, endemism, and AOH density

3 Results

3.1 Accuracy assessment

3.1.1 Species-level validation

Following the validation framework proposed by Dahal et al. (2021), we assessed the reliability of species-level AOH predictions using two complementary approaches: a continuous-suitability evaluation and a threshold-based binary evaluation. Dahal et al. developed this framework for AOH validation under presence-only data conditions, using point localities to test whether predicted suitable areas perform better than random expectation within a species' range. In that framework, model prevalence is defined as the proportion of a species' range retained as suitable habitat, and point prevalence as the proportion of independent point localities falling within the predicted suitable area. AOH predictions are considered better than random when point prevalence exceeds model prevalence.

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

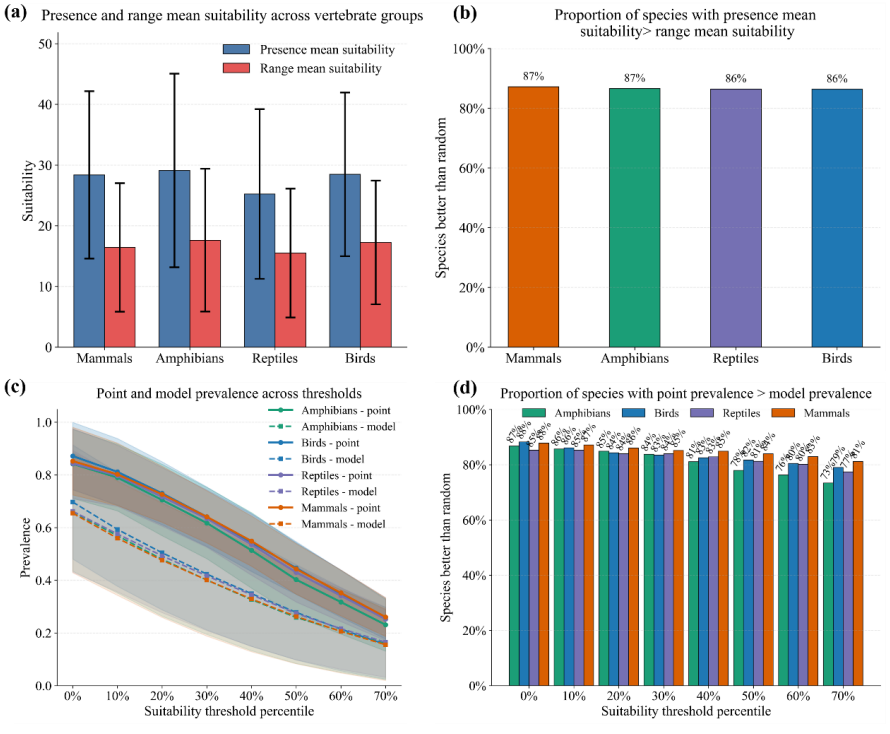
For the continuous evaluation, we compared the mean predicted suitability extracted at independent occurrence points with the mean predicted suitability across the broader evaluation domain for each species. Across all four terrestrial vertebrate groups, occurrence points were consistently associated with higher suitability values than the range-wide average (Fig. 7a). This pattern indicates that the continuous AOH surfaces successfully concentrated higher suitability in locations where species were actually observed, rather than distributing suitability uniformly across the entire range. The same signal was evident at the species level: the proportion of species with presence mean suitability exceeding range mean suitability reached 87 % for mammals, 87 % for amphibians, 86 % for reptiles, and 86 % for birds (Fig. 7b). Thus, for the great majority of species, observed presences were located in areas predicted to be more suitable than expected from the range background, demonstrating that the continuous predictions captured meaningful within-range spatial variation.

We further evaluated binary AOH predictions across a gradient of suitability-threshold percentiles by comparing point prevalence with model prevalence, following the logic of Dahal et al. (2021). In our results, point prevalence remained consistently higher than model prevalence across all taxa and across all tested thresholds (Fig. 7c). This indicates that the thresholded binary AOH maps retained independent occurrence points more effectively than expected from retained area alone, confirming that predicted suitable habitat was non-randomly concentrated in ecologically relevant portions of species ranges. As expected, both point prevalence and model prevalence declined progressively with increasing suitability thresholds, reflecting the more restrictive nature of conservative binary AOH definitions. However, the performance advantage of point prevalence over model prevalence persisted throughout the threshold gradient. At the species level, the proportion of species with point prevalence exceeding model prevalence was high at low thresholds and declined only gradually under stricter thresholds (Fig. 7d). At the 0th percentile threshold, this proportion was

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

87 % for amphibians, 88 % for birds, 87 % for reptiles, and 88 % for mammals; even at the 70th percentile threshold, it remained 73 %, 79 %, 77 %, and 81 %, respectively. Mammals consistently showed the highest proportions across thresholds, whereas amphibians exhibited the strongest decline at higher thresholds, suggesting somewhat greater sensitivity to increasingly restrictive suitability cutoffs.

Overall, the species-level validation provides strong support for the reliability of the predicted AOH maps. Under the continuous evaluation, independent occurrence records were overwhelmingly associated with higher predicted suitability than the broader range background. Under the binary evaluation, thresholded AOH maps consistently retained occurrence points at rates exceeding background expectation. Together, these results show that the predicted species distributions are not only spatially structured in an ecologically meaningful way, but also robust across alternative validation formulations based on the standard proposed by Dahal et al. (2021).



https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

Fig. 5 Species-level validation of species distribution predictions across four terrestrial vertebrate taxa. *Panels a,b present the continuous-model validation, in which presence mean suitability extracted at occurrence points was compared with range mean suitability calculated across the full species range and evaluation domain. Panel a shows taxon-specific means, and panel b shows the proportion of species with presence mean suitability greater than range mean suitability. Panels c,d present the binary AOH validation, in which continuous suitability predictions were converted to binary AOH maps using different suitability-threshold percentiles. Panel c compares point prevalence and model prevalence across thresholds, where point prevalence is the proportion of occurrence points falling within the binary AOH and model prevalence is the proportion of the species range retained as AOH. Panel d shows the proportion of species for which point prevalence exceeded model prevalence, indicating binary AOH predictions that performed better than random.*

3.1.2 Dataset-level validation

To evaluate the spatial consistency of the dataset at broader spatial scales, we compared AOH-derived estimated richness with occurrence-derived observed richness across terrestrial ecoregions (Fig. 6). For each ecoregion, observed richness was calculated as the number of unique species represented by independent occurrence records, whereas estimated richness was obtained by aggregating species-level AOH predictions within the same ecoregion. This comparison was used to assess whether the dataset could reproduce major geographic gradients and inter-ecoregion variation in biodiversity richness.

At the global scale, the estimated richness map reproduced the main spatial structure observed in occurrence-derived richness (Fig. 6a, c). Both datasets showed clear richness hotspots in tropical regions, especially in the Amazon Basin, Central Africa, and South and Southeast Asia, and substantially lower richness in arid regions, high latitudes, and other environmentally harsh areas. The latitudinal richness profiles

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

498 were also highly consistent between the two datasets (Fig. 6b, d; Fig. S7), with both
499 estimated and observed richness peaking in the tropics and declining toward the poles.
500 This agreement indicates that the AOH-based dataset successfully captured the major
501 global biodiversity gradients reflected by independent occurrence data.

502 The correspondence between estimated and observed richness was also strong at
503 the ecoregion level. Across all terrestrial vertebrates combined, estimated richness was
504 highly correlated with observed richness among ecoregions (Pearson's r=0.940; Fig.
505 6e). Similar positive relationships were obtained for each taxonomic group separately,
506 including mammals (r=0.785; Fig. S3), amphibians (r=0.920; Fig. S4), reptiles (r=0.898;
507 Fig. S5), and birds (r=0.953; Fig. S6). These results demonstrate that the dataset
508 retained substantial ecological realism not only for total vertebrate richness but also
509 within individual taxonomic groups.

510 Although the overall agreement was high, deviations from the 1:1 relationship
511 were not random. For all terrestrial vertebrates combined, estimated richness was lower
512 than observed richness in 57.0 % of ecoregions and higher than observed richness in
513 43.0 % (Fig. 6f). However, these two categories differed markedly in their sampling
514 intensity. Ecoregions where estimated richness exceeded observed richness generally
515 had substantially lower sampling intensity than those where observed richness equaled
516 or exceeded estimated richness (Fig. 6e, g). This pattern suggests that many cases of
517 estimated richness being higher than observed richness are likely attributable to
518 incomplete occurrence sampling rather than systematic overprediction by the model. In
519 poorly sampled ecoregions, occurrence-derived richness is expected to underestimate
520 the true number of species present, whereas the AOH-based estimates can recover part
521 of this missing diversity by integrating habitat and climatic suitability. This
522 interpretation was further supported by taxon-specific analyses. For mammals, reptiles,
523 and amphibians, estimated richness exceeded observed richness in 81.5 %, 69.7 %, and
524 66.9 % of ecoregions, respectively (Figs. S3–S5), and these positive deviations were

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026


Open Access Earth System Science Data Discussions

525 again concentrated in regions with relatively low sampling intensity. In contrast, birds
526 showed the opposite tendency: observed richness was higher than estimated richness in
527 80.4 % of ecoregions (Fig. S6), likely reflecting the much denser and more spatially
528 extensive occurrence coverage available for birds (Fig. S1-S2).

529 Overall, the dataset-level validation demonstrates that the AOH-based richness
530 estimates reliably reproduce the global spatial pattern of terrestrial vertebrate richness.
531 The strong ecoregion-level correlations across all taxa indicate that the dataset captures
532 broad-scale biodiversity variation well. At the same time, the relationship between
533 richness discrepancies and sampling intensity shows that occurrence-derived richness
534 is itself sensitive to uneven sampling effort, especially in poorly surveyed ecoregions.
535 Under these conditions, the AOH-based estimates likely provide a more complete
536 representation of true biodiversity patterns than occurrence records alone.

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

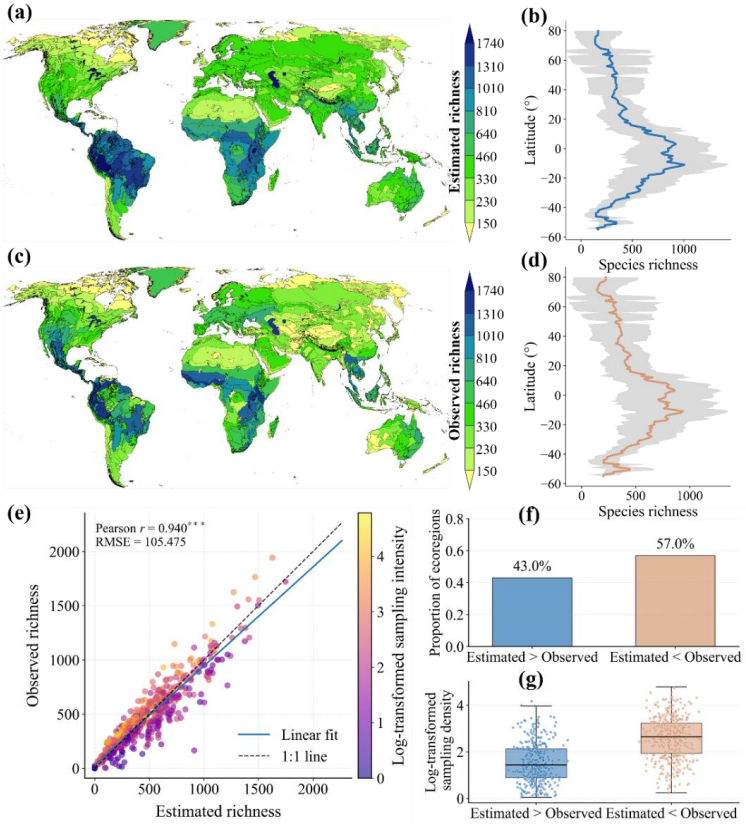


Fig. 6 Dataset-level validation of global estimated species richness against occurrence-derived richness across terrestrial ecoregions. *Panel a shows the spatial distribution of estimated species richness in 2020 derived from this study, and panel c shows the corresponding observed species richness calculated from independent occurrence records. Panels b and d summarize the latitudinal richness profiles, where each value represents the mean species richness of all ecoregions intersecting a given latitude. Panel e compares estimated and observed richness across ecoregions, with each point representing one ecoregion and point color indicating the mean sampling intensity within that ecoregion (Fig. S2). Panel f shows the proportion of ecoregions in which estimated richness is greater than or lower than observed richness, and panel g compares the distribution of sampling intensity between these two categories.*

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

3.2 Global patterns of ecoregion-level biodiversity indicators

In addition to species richness (Fig. 6), the other three biodiversity indicators—threatened species richness, species endemism, and AOH density—also showed clear and spatially coherent global patterns across terrestrial ecoregions in 2020 (Fig. 7). All three indicators were generally highest in tropical regions and lowest in deserts, high-latitude regions, and other environmentally harsh areas. Threatened species richness showed pronounced hotspots in Central and South America, sub-Saharan Africa, South and Southeast Asia, whereas species endemism was more strongly concentrated in parts of the tropical Andes, Mesoamerica, eastern Africa, Madagascar, and Southeast Asia. AOH density broadly followed the global pattern of overall biodiversity, with highest values in humid tropical ecoregions. The corresponding latitudinal profiles all showed clear tropical peaks and declining values toward higher latitudes.

Biome-level comparisons further revealed clear differences among the eight biome groups (Fig. 8). Tropical forests consistently showed the highest values for all four biodiversity indicators and also exhibited the greatest among-ecoregion variation. Temperate forests and tropical grasslands generally occupied intermediate positions, although some ecoregions reached relatively high values. Temperate grasslands, Mediterranean, and deserts & xeric shrublands tended to show lower values overall, whereas boreal forests and tundra generally had the lowest values across indicators.

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

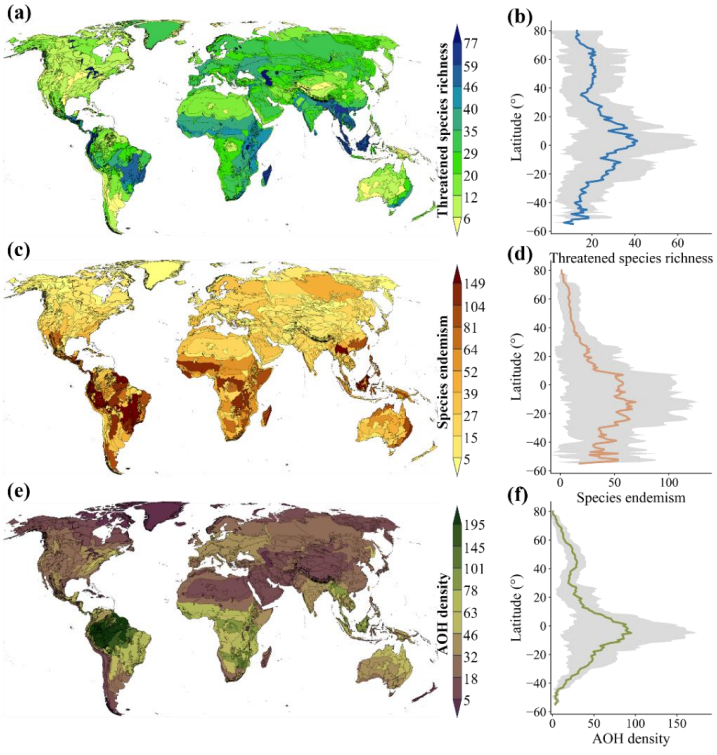


Fig. 7 Global patterns of estimated threatened species richness, species endemism, and AOH density across terrestrial ecoregions in 2020. *Panels a, c, and e show the spatial distributions of estimated threatened species richness, species endemism, and AOH density in 2020, respectively. Panels b, d, and f show the corresponding latitudinal profiles, where each value represents the mean value of all ecoregions intersecting a given latitude.*

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

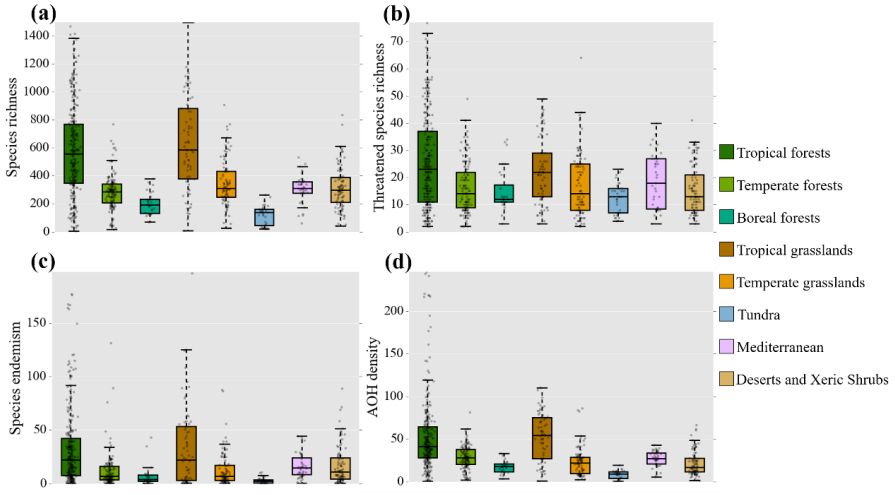


Fig. 8 Distribution of four estimated biodiversity indicators across terrestrial ecoregions within different biomes in 2020. *Panels a–d show the biome-level distributions of species richness, threatened species richness, species endemism, and AOH density, respectively. In each panel, each point represents a single ecoregion. The box indicates the interquartile range (IQR, 25th–75th percentiles), the horizontal line inside the box marks the median, and the whiskers extend to the range of values within 1.5 × IQR.*

3.3 Projected changes in ecoregion-level biodiversity indicators

Projected changes in biodiversity indicators from 2020 to 2100 showed pronounced spatial heterogeneity under both SSP245 and SSP585 (Fig. 9). In general, species richness, threatened species richness, and AOH density tended to decline across many tropical and subtropical ecoregions, particularly in parts of Central and South America, Africa, and South and Southeast Asia, whereas increases were more often projected in some temperate and high-latitude regions. Species endemism showed a more spatially mixed pattern, but declines were still concentrated in several biodiversity-rich tropical areas. Across all four indicators, the magnitude and extent of change were generally greater under SSP585 than under SSP245, indicating that more

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

593 severe climate and habitat change would lead to stronger biodiversity redistribution and
594 loss. The frequency distributions and latitudinal profiles further highlighted this
595 contrast between low and high latitudes (Fig. 9). Negative changes were concentrated
596 mainly in tropical regions, while positive changes occurred more frequently toward
597 temperate and boreal zones. This pattern was especially evident for species richness
598 and AOH density, both of which showed clear tropical losses and poleward gains.
599 Threatened species richness was more consistently negative across latitudes, especially
600 under SSP585, suggesting a broad reduction in the concentration of species of
601 conservation concern. By contrast, species endemism changed less uniformly,
602 reflecting the more localized nature of range restriction.

603 Biome-level comparisons also revealed clear differences among the eight biome
604 groups (Fig. 10). Under both scenarios, tropical forests showed the largest median
605 declines and the greatest variation among ecoregions for most indicators, particularly
606 species richness and AOH density. Tropical grasslands also experienced notable
607 declines, whereas temperate forests and temperate grasslands generally showed smaller
608 changes. Boreal forests and tundra more often exhibited small losses or slight increases,
609 especially for species richness. Overall, projected changes were consistently more
610 negative under SSP585 than under SSP245 across most biomes and indicators,
611 reinforcing the greater biodiversity risk under the higher-emissions scenario.

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

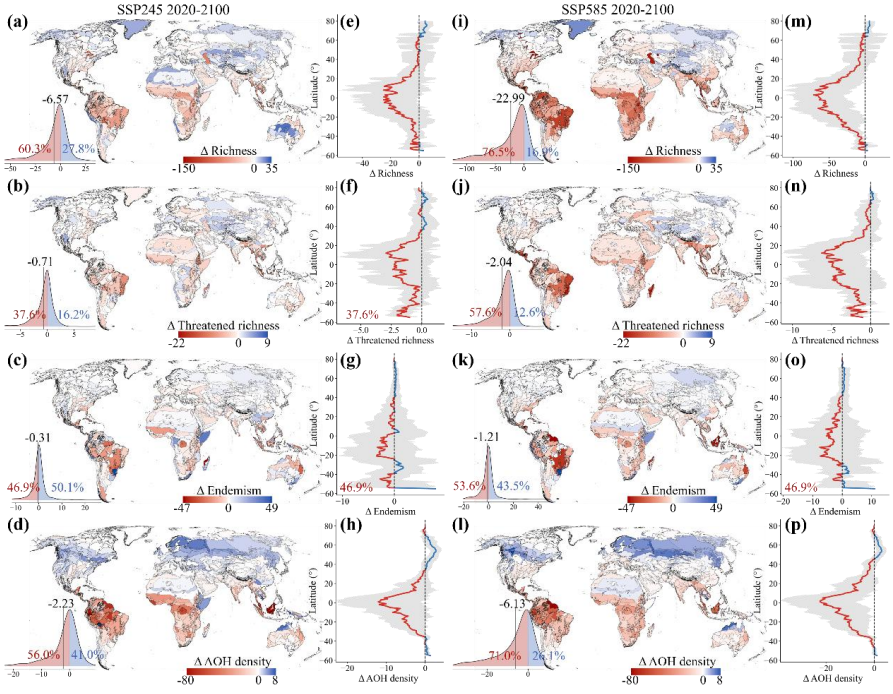
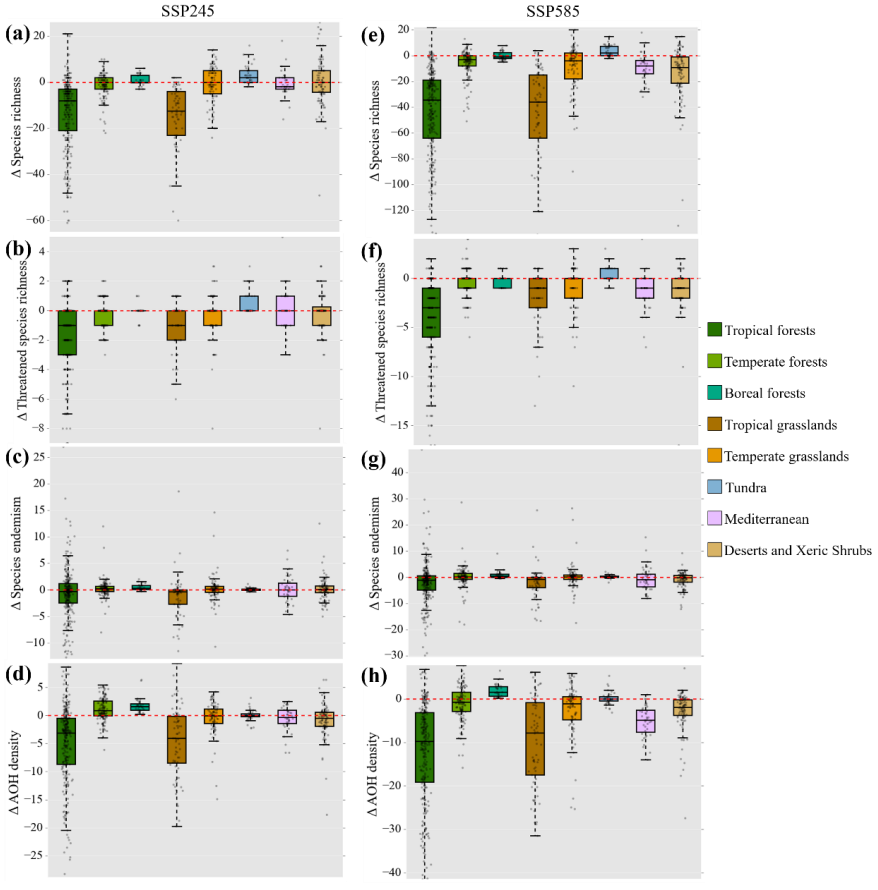


Fig. 9 Projected changes in biodiversity indicators across terrestrial ecoregions from 2020 to 2100 under SSP245 and SSP585. scenarios. *Panels a–d show projected changes in species richness, threatened species richness, species endemism, and AOH density, respectively, under the SSP245 scenario, while panels i–l show the corresponding changes under SSP585. In each map, red colors indicate decreases and blue colors indicate increases relative to 2020. The inset frequency distributions summarize changes across all ecoregions; the black vertical line marks the mean change across all ecoregions. Panels e–h (SSP245) and m–p (SSP585) show the corresponding latitudinal profiles of projected change, where each value represents the mean change of all ecoregions intersecting a given latitude.*

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026



623

624 ***Fig. 10 Biome-level distributions of projected changes in biodiversity indicators***
625 ***across terrestrial ecoregions under SSP245 and SSP585.*** *Panels a–d show the*
626 *distributions of projected changes in species richness, threatened species richness,*
627 *species endemism, and AOH density under SSP245, respectively, while panels e–h*
628 *show the corresponding distributions under SSP585. Each point represents one*
629 *ecoregion. Boxplots summarize the median, interquartile range, and whiskers*
630 *extending to 1.5 × IQR for all ecoregions within each biome. The red dashed line*
631 *indicates no change relative to 2020.*

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

4 Discussion

The dataset presented here provides a new basis for analyzing biodiversity change at an ecologically meaningful spatial unit by integrating climate suitability, habitat suitability, and species range constraints into long-term ecoregion-level indicators for terrestrial vertebrates. This is important because both climate change and land-use change are widely recognized as major direct drivers of biodiversity loss, yet global assessments still often rely either on coarse expert range maps or on occurrence records that are uneven in space and time (Newbold et al., 2015; Pereira et al., 2010). By using the Area of Habitat (AOH) framework, our dataset helps bridge this gap: AOH has been proposed as a more ecologically realistic representation of species distributions because it refines broad ranges by excluding unsuitable land cover and elevation, and recent global applications have shown its utility for biodiversity assessment and conservation planning (Chapman et al., 2025; Brooks et al., 2019; Hurlbert and Jetz, 2007). Moreover, the use of ecoregions as the reporting unit makes the dataset particularly relevant for macroecological comparison and regional conservation prioritization, because ecoregions better reflect biogeographic structure and species turnover than administrative units or regular grids (Olson et al., 2001; Dinerstein et al., 2017).

Beyond documenting broad spatial patterns, our results show why a multidimensional indicator framework is necessary for biodiversity assessment. Species richness alone captures the number of species, but it cannot distinguish between areas containing many widespread species and those containing a high concentration of threatened or range-restricted taxa (Jetz et al., 2019; Watermeyer et al., 2021). In our dataset, threatened richness, endemism, and AOH density revealed spatial structures that only partly overlapped with total richness, indicating that different ecoregions contribute different dimensions of biodiversity value. This is consistent with previous work showing that alternative biodiversity metrics often identify different conservation priorities and should not be treated as interchangeable (Jetz et al., 2019). The broader

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

659 biodiversity-monitoring literature has likewise emphasized that model-based, spatially
660 continuous indicators are essential for translating heterogeneous species data into
661 policy-relevant measures of status and change across space and time.

662 The projected declines concentrated in many tropical ecoregions are also
663 ecologically and conservation-relevant. Tropical regions have long been recognized as
664 areas where exceptional concentrations of endemic species coincide with severe habitat
665 loss and strong conservation pressure (Brooks et al., 2002; Myers et al., 2000). Our
666 finding that tropical forests and tropical grasslands tend to experience larger declines
667 than most higher-latitude biomes is therefore consistent with the broader view that
668 biodiversity-rich regions are especially vulnerable to ongoing climate and land-use
669 change. This point is particularly important for endemism: endemic species are
670 generally more exposed to climate-related risk because their restricted distributions
671 leave less scope for spatial buffering (Manes et al., 2021). Accordingly, apparent gains
672 in some temperate or boreal ecoregions should not be interpreted as offsetting losses in
673 tropical systems, because spatial redistribution of species does not compensate for
674 declines in irreplaceable, range-restricted, or threatened biotas (Daru et al., 2021). The
675 stronger and more widespread changes under SSP585 are in line with previous global
676 assessments showing that biodiversity responses become increasingly severe under
677 higher-emissions pathways and when climate change interacts with land-use change
678 (Newbold, 2018; Di Marco et al., 2019). At the same time, the stronger declines in
679 threatened richness than in total richness in many regions suggest that conservation risk
680 may intensify even where overall richness changes appear moderate. This reinforces
681 the need to evaluate future biodiversity change through multiple indicators rather than
682 richness alone, especially in conservation planning contexts where the relevant
683 objective may be to retain threatened taxa, endemic taxa, or habitat-supported
684 persistence rather than simply maintain species counts.

685 Several limitations should nevertheless be recognized when interpreting this

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

686 dataset. First, although the AOH framework improves ecological realism relative to
687 coarse range maps, the resulting indicators still inherit uncertainty from underlying
688 species ranges, habitat coding, occurrence coverage, and climate projections. Second,
689 like most large-scale distribution models, our projections do not explicitly incorporate
690 dispersal limitation, demographic feedbacks, biotic interactions, evolutionary
691 adaptation, or time lags in species responses (Urban et al., 2016; Wisz et al., 2013;
692 Dormann et al., 2018). These omissions matter because future range shifts and
693 community reassembly are shaped not only by abiotic suitability, but also by processes
694 that may either constrain or facilitate persistence under environmental change. Recent
695 modelling studies have further shown that uncertainty in future species projections can
696 increase strongly through time and, in some cases, may exceed the uncertainty arising
697 from divergent climate models themselves (Thuiller et al., 2019). These issues do not
698 diminish the value of the present dataset, but they do indicate that projected changes
699 should be interpreted as scenario-based estimates of potential redistribution rather than
700 precise forecasts of realized future assemblages.

701 The limitations in this study also point to several promising directions for future
702 work. A major next step will be to extend the framework beyond terrestrial vertebrates
703 to plants, invertebrates, and freshwater-dependent taxa, which would allow a more
704 complete characterization of ecoregion-level biodiversity change. It would also be
705 valuable to integrate explicit uncertainty layers, additional scenario ensembles, and
706 process-based components related to dispersal, demographic persistence, and species
707 interactions. More broadly, this type of long-term, model-integrated dataset aligns well
708 with current efforts to build essential biodiversity variables and other harmonized
709 monitoring products that link raw species observations, environmental covariates, and
710 decision-ready indicators (Jetz et al., 2019; Pereira et al., 2013). By providing spatially
711 explicit estimates across historical and future periods, the dataset can support
712 biodiversity monitoring, scenario comparison, conservation prioritization, and the

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

evaluation of global targets under rapidly changing environmental conditions.

5 Data availability

The data for this manuscript are available at https://doi.org/10.5281/zenodo.20119261 (Zheng, 2026).

6 Conclusions

This study developed a global long-term ecoregion-level dataset of biodiversity indicators for terrestrial vertebrates under climate and habitat change. By integrating species occurrence records, expert-derived range maps, climate data, and temporally explicit habitat maps within an AOH-based framework, we generated harmonized estimates of species richness, threatened species richness, species endemism, and AOH density across global terrestrial ecoregions. Validation results showed that the dataset is reliable at both the species and dataset levels. Predicted AOH performed well in identifying independent occurrence locations, and estimated species richness reproduced the major spatial patterns of occurrence-derived species richness across ecoregions. The dataset further revealed strong tropical concentration of biodiversity indicators and projected widespread future declines in many tropical ecoregions, especially under SSP585. Overall, this dataset provides a globally consistent and ecologically meaningful resource for biodiversity monitoring, macroecological analysis, and conservation assessment. It offers a useful basis for tracking long-term biodiversity change across terrestrial ecoregions under ongoing climate and habitat change.

Author contributions

Conceptualization: SZ, HW, LY, KM; methodology: SZ, HW; software: SZ; results validation: SZ; analysis: SZ, HW, LY; data collection: SZ, HW; writing: SZ; visualization: SZ; supervision: LY; funding acquisition: LY; project administration: LY.

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

Competing interests

The contact author has declared that none of the authors have any competing interests.

Acknowledgements

This research was funded by the National Key R&D Program of China (2024YFF1307600) and the National Natural Science Foundation of China (42401314).

References

Beck, J., Böller, M., Erhardt, A., and Schwanghart, W.: Spatial bias in the GBIF database and its effect on modeling species' geographic distributions, Ecological Informatics, 19, 10–15, 2014.

BirdLife, I.: Bird species distribution maps of the world, BirdLife International, Cambridge, UK and NatureServe, Arlington, USA, 2015.

Brooks, T. M., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., Rylands, A. B., Konstant, W. R., Flick, P., Pilgrim, J., Oldfield, S., and Magin, G.: Habitat loss and extinction in the hotspots of biodiversity, Conservation biology, 16, 909–923, 2002.

Brooks, T. M., Pimm, S. L., Akçakaya, H. R., Buchanan, G. M., Butchart, S. H., Foden, W., Hilton-Taylor, C., Hoffmann, M., Jenkins, C. N., and Joppa, L.: Measuring terrestrial area of habitat (AOH) and its utility for the IUCN Red List, Trends in ecology & evolution, 34, 977–986, 2019.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., and Wardle, D. A.: Biodiversity loss and its impact on humanity, Nature, 486, 59–67, 2012.

Change, I. C.: The physical science basis, (No Title), 2013.

Chapman, M., Jung, M., Leclère, D., Boettiger, C., Augustynczik, A. L., Gusti, M., Ringwald, L., and Visconti, P.: Meeting European Union biodiversity targets under future land-use demands, Nature Ecology & Evolution, 9, 810–821, 2025.

Chen, G., Li, X., and Liu, X.: Global land projection based on plant functional types with a 1-km resolution under socio-climatic scenarios, Scientific Data, 9, 125, 2022.

Dahal, P. R., Lumbierres, M., Butchart, S. H., Donald, P. F., and Rondinini, C.: A validation standard for area of habitat maps for terrestrial birds and mammals, Geoscientific Model Development Discussions, 2021, 1–25, 2021.

Daru, B. H., Davies, T. J., Willis, C. G., Meineke, E. K., Ronk, A., Zobel, M., Pärtel, M., Antonelli, A., and Davis, C. C.: Widespread homogenization of plant communities in the Anthropocene, Nature communications, 12, 6983, 2021.

Di Marco, M., Harwood, T. D., Hoskins, A. J., Ware, C., Hill, S. L., and Ferrier,

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

773 S.: Projecting impacts of global climate and land-use scenarios on plant biodiversity
774 using compositional-turnover modelling, Global change biology, 25, 2763–2778, 2019.
775 Díaz, S., Settele, J., Brondízio, E. S., Ngo, H. T., Agard, J., Arneth, A., Balvanera,
776 P., Brauman, K. A., Butchart, S. H., and Chan, K. M.: Pervasive human-driven decline
777 of life on Earth points to the need for transformative change, Science, 366, eaax3100,
778 2019.
779 Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake,
780 E., Hahn, N., Palminteri, S., Hedao, P., and Noss, R.: An ecoregion-based approach to
781 protecting half the terrestrial realm, BioScience, 67, 534–545, 2017.
782 Dormann, C. F., Bobrowski, M., Dehling, D. M., Harris, D. J., Hartig, F., Lischke,
783 H., Moretti, M. D., Pagel, J., Pinkert, S., and Schleuning, M.: Biotic interactions in
784 species distribution modelling: 10 questions to guide interpretation and avoid false
785 conclusions, Global ecology and biogeography, 27, 1004–1016, 2018.
786 Elith, J. and Leathwick, J. R.: Species distribution models: ecological explanation
787 and prediction across space and time, Annual review of ecology, evolution, and
788 systematics, 40, 677–697, 2009.
789 Franklin, J.: Mapping Species Distributions, American history, 1861, 1945.
790 Gearty, W. and Chamberlain, S.: rredlist:'IUCN'Red List client, R package version
791 0.7, 1, 2022.
792 Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H.,
793 Lentini, P. E., McCarthy, M. A., Tingley, R., and Wintle, B. A.: Is my species
794 distribution model fit for purpose? Matching data and models to applications, Global
795 ecology and biogeography, 24, 276–292, 2015.
796 Hooper, D. U., Adair, E. C., Cardinale, B. J., Byrnes, J. E., Hungate, B. A.,
797 Matulich, K. L., Gonzalez, A., Duffy, J. E., Gamfeldt, L., and O'Connor, M. I.: A global
798 synthesis reveals biodiversity loss as a major driver of ecosystem change, Nature, 486,
799 105–108, 2012.
800 Hurlbert, A. H. and Jetz, W.: Species richness, hotspots, and the scale dependence
801 of range maps in ecology and conservation, Proceedings of the National Academy of
802 Sciences, 104, 13384–13389, 2007.
803 IPBES, W.: Intergovernmental science-policy platform on biodiversity and
804 ecosystem services, Summary for policy makers of the global assessment report on
805 biodiversity and ecosystem services of the intergovernmental science-policy platform
806 on biodiversity and ecosystem services. IPBES Secretariat, Bonn, Germany, 2019.
807 Isbell, F., Gonzalez, A., Loreau, M., Cowles, J., Díaz, S., Hector, A., Mace, G. M.,
808 Wardle, D. A., O'Connor, M. I., and Duffy, J. E.: Linking the influence and dependence
809 of people on biodiversity across scales, Nature, 546, 65–72, 2017.
810 Jetz, W., McGeoch, M. A., Guralnick, R., Ferrier, S., Beck, J., Costello, M. J.,
811 Fernandez, M., Geller, G. N., Keil, P., and Merow, C.: Essential biodiversity variables
812 for mapping and monitoring species populations, Nature ecology & evolution, 3, 539–
813 551, 2019.
814 Jung, M., Dahal, P. R., Butchart, S. H., Donald, P. F., De Lamo, X., Lesiv, M.,

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

815 Kapos, V., Rondinini, C., and Visconti, P.: A global map of terrestrial habitat types,
816 Scientific data, 7, 256, 2020.
817 Lobo, J. M. and Tognelli, M. F.: Exploring the effects of quantity and location of
818 pseudo-absences and sampling biases on the performance of distribution models with
819 limited point occurrence data, Journal for Nature Conservation, 19, 1–7, 2011.
820 Lumbierres, M., Dahal, P. R., Di Marco, M., Butchart, S. H., Donald, P. F., and
821 Rondinini, C.: Translating habitat class to land cover to map area of habitat of terrestrial
822 vertebrates, Conservation Biology, 36, e13851, 2022.
823 Maiorano, L.: Global habitat suitability models of terrestrial mammals, ⋯ of the
824 Royal⋯, 2011.
825 Manes, S., Costello, M. J., Beckett, H., Debnath, A., Devenish-Nelson, E., Grey,
826 K.-A., Jenkins, R., Khan, T. M., Kiessling, W., and Krause, C.: Endemism increases
827 species' climate change risk in areas of global biodiversity importance, Biological
828 conservation, 257, 109070, 2021.
829 Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S.,
830 Caud, N., Chen, Y., Goldfarb, L., and Gomis, M. I.: Climate change 2021: the physical
831 science basis, Contribution of working group I to the sixth assessment report of the
832 intergovernmental panel on climate change, 2, 2391, 2021.
833 Meyer, C., Weigelt, P., and Kreft, H.: Multidimensional biases, gaps and
834 uncertainties in global plant occurrence information, Ecology letters, 19, 992–1006,
835 2016.
836 Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., and Kent,
837 J.: Biodiversity hotspots for conservation priorities, Nature, 403, 853–858, 2000.
838 Newbold, T.: Future effects of climate and land-use change on terrestrial vertebrate
839 community diversity under different scenarios, Proceedings of the Royal Society B:
840 Biological Sciences, 285, 2018.
841 Newbold, T., Hudson, L. N., Hill, S. L., Contu, S., Lysenko, I., Senior, R. A.,
842 Börger, L., Bennett, D. J., Choimes, A., and Collen, B.: Global effects of land use on
843 local terrestrial biodiversity, Nature, 520, 45–50, 2015.
844 Olson, D., Dinerstein, E., Wikramanayake, E., Burgess, N., Powell, G.,
845 Underwood, E., d'Amico, J., Itoua, I., Strand, H., and Morrison, J.: Terrestrial
846 ecoregions of the world: a new map of life on earth Bioscience 51: 933–938, 2001.
847 Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H., Scholes, R.
848 J., Bruford, M. W., Brummitt, N., Butchart, S. H., and Cardoso, A.: Essential
849 biodiversity variables, Science, 339, 277–278, 2013.
850 Pereira, H. M., Leadley, P. W., Proença, V., Alkemade, R., Scharlemann, J. P.,
851 Fernandez-Manjarrés, J. F., Araújo, M. B., Balvanera, P., Biggs, R., and Cheung, W.
852 W.: Scenarios for global biodiversity in the 21st century, Science, 330, 1496–1501,
853 2010.
854 Pettorelli, N., Laurance, W. F., O'Brien, T. G., Wegmann, M., Nagendra, H., and
855 Turner, W.: Satellite remote sensing for applied ecologists: opportunities and challenges,

https://doi.org/10.5194/essd-2026-354
Preprint. Discussion started: 29 June 2026

856 Journal of Applied Ecology, 51, 839–848, 2014.
857 Phillips, S. J. and Dudík, M.: Modeling of species distributions with Maxent: new
858 extensions and a comprehensive evaluation, Ecography, 31, 161–175, 2008.
859 Santini, L., Benítez-López, A., Maiorano, L., Čengić, M., and Huijbregts, M. A.:
860 Assessing the reliability of species distribution projections in climate change research,
861 Diversity and Distributions, 27, 1035–1050, 2021.
862 Secretariat, G.: GBIF: Global Biodiversity Information Facility, 2020.
863 Smith, J. R., Beaury, E. M., Cook-Patton, S. C., and Levine, J. M.: Variable
864 impacts of land-based climate mitigation on habitat area for vertebrate diversity,
865 Science, 387, 420–425, 2025.
866 Thuiller, W., Guéguen, M., Renaud, J., Karger, D. N., and Zimmermann, N. E.:
867 Uncertainty in ensembles of global biodiversity scenarios, Nature communications, 10,
868 1446, 2019.
869 Tilman, D., Clark, M., Williams, D. R., Kimmel, K., Polasky, S., and Packer, C.:
870 Future threats to biodiversity and pathways to their prevention, Nature, 546, 73–81,
871 2017.
872 Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J.-B., Pe'er, G., Singer, A., Bridle,
873 J. R., Crozier, L. G., De Meester, L., and Godsoe, W.: Improving the forecast for
874 biodiversity under climate change, Science, 353, aad8466, 2016.
875 Watermeyer, K. E., Guillera-Arroita, G., Bal, P., Burgass, M. J., Bland, L. M.,
876 Collen, B., Hallam, C., Kelly, L. T., McCarthy, M. A., and Regan, T. J.: Using decision
877 science to evaluate global biodiversity indices, Conservation Biology, 35, 492–501,
878 2021.
879 Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F.,
880 Dormann, C. F., Forchhammer, M. C., Grytnes, J. A., and Guisan, A.: The role of biotic
881 interactions in shaping distributions and realised assemblages of species: implications
882 for species distribution modelling, Biological reviews, 88, 15–30, 2013.
883 Yu, L., Du, Z., Li, X., Gu, J., Li, X., Zhong, L., Duojiweise, Wu, H., Zhao, Q., and
884 Ma, X.: FROM-GLC Plus 3.0: Multimodal Land Change Mapping with SAM and
885 Dense Surface Observations, Journal of Remote Sensing, 5, 0728, 2025.
886 Zheng, S., Yu, L., & Wu, H: Global dynamic maps of terrestrial Area-of-Habitat
887 (AOH) [dataset], https://doi.org/10.5281/zenodo.20119261, 2026.
888 Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D.,
889 Farooq, H., Herdean, A., Ariza, M., and Scharn, R.: CoordinateCleaner: Standardized
890 cleaning of occurrence records from biological collection databases, Methods in
891 Ecology and Evolution, 10, 744–751, 2019.

892