Supplementary Materials for

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

Shijun Zheng et al.

*Corresponding author. Email: leyu@tsinghua.edu.cn

Supplemental Text

Text S1 Reconstruction of temporally explicit habitat maps

Historical habitat reconstruction. To extend the 2015 global terrestrial habitat map into a temporally continuous dataset suitable for long-term biodiversity modelling, we reconstructed habitat composition for both historical and future periods using time-varying land-cover and vegetation datasets. The 2015 habitat map of Jung et al. (2020), which is directly aligned with the IUCN habitat classification system, was used as the baseline reference. Historical habitat changes were reconstructed for 1990–2020 using the From-GLC Plus land-cover dataset (Yu et al., 2025), whereas future habitat composition for 2030, 2050, and 2100 under SSP245 and SSP585 was projected using the plant functional type (PFT) dataset of Chen et al. (2022). For the historical period, we used annual land-cover layers from From-GLC Plus to estimate changes in broad habitat composition relative to the 2015 baseline. The original From-GLC Plus data were provided at 30 m spatial resolution and include 10 Level-1 land-cover classes. To ensure consistency with the habitat classification framework adopted in this study, these land-cover classes were aggregated into four broad habitat categories: (1) forest, including plantations; (2) non-forest natural vegetation; (3) agriculture, including pasture and rural gardens; and (4) urban areas (Table S1). For each year from 1990 to 2020, we compared the corresponding annual land-cover layer with the 2015 layer and calculated the proportional differences in these four habitat categories for each pixel. The resulting annual change layers were then aggregated to 10 km resolution, such that each grid cell represented the net change in the proportional coverage of the four habitat categories relative to 2015. These annual change estimates were then used to reconstruct habitat-type composition within each 10 km grid cell. For urban areas, proportional changes were applied directly to the urban habitat class. For the other three broad categories, each of which contains multiple IUCN Level 2 habitat types, gains or losses were distributed among constituent habitat types according to their relative proportions in the 2015 baseline map. This procedure preserved the internal composition of each broad habitat category while allowing total category area to vary through time. In some grid cells, a broad habitat category was newly introduced in a given year despite being absent in the 2015 baseline map. This situation occurred mainly where agricultural expansion or urban development created habitat classes not previously represented within the cell. In such cases, the composition of the newly introduced category was assigned according to the mean proportional composition of that category within the corresponding ecoregion. If proportional adjustments caused any habitat type to become negative, its value was set to zero. After all adjustments, habitat proportions within each grid cell were rescaled so that the sum of all habitat classes equaled one. This procedure produced annual habitat raster stacks for 1990–2020, each containing 58 IUCN-compatible Level 2 habitat layers.

Future habitat reconstruction. Future habitat composition was projected using the global PFT dataset of Chen et al. (2022), which provides spatially explicit proportional cover of 17 plant functional types at 5-year intervals under SSP245 and SSP585. We used the 2015 PFT layer as the reference and extracted projected layers for 2030, 2050, and 2100 under both scenarios. The original data were provided at 1 km spatial resolution. To maintain consistency with the historical reconstruction and the habitat classification applied throughout the study, the 17 plant functional types were aggregated into the same four broad habitat categories used above: forest, non-forest natural vegetation, agriculture, and urban areas (Table S1). For each projection year and scenario, we calculated the differences in the proportional coverage of these four categories relative to 2015 at the pixel level, and then aggregated the results to 10 km resolution. These projected category-level changes were subsequently translated into changes in the proportional cover of the 58 Level 2 habitat types within each 10 km grid cell. As in the historical reconstruction, changes in urban area were applied directly to the urban habitat class, whereas changes in forest, non-forest natural vegetation, and agriculture were allocated among their constituent habitat types according to their baseline proportions in 2015. Where a habitat category newly appeared in a grid cell, its internal habitat-type composition was estimated using the mean composition of that category within the same ecoregion. Negative values generated during the adjustment process were truncated to zero, and all habitat proportions were rescaled so that the total across habitat types summed to one for each grid cell. This procedure generated six future habitat raster stacks representing IUCN-compatible habitat composition for 2030, 2050, and 2100 under SSP245 and SSP585.

Text S2 Sampling-bias correction and generation of background points for climatic suitability modelling

To account for strong spatial bias in global occurrence records, we implemented a taxon-specific background sampling strategy before fitting the MaxEnt models. This procedure was designed to ensure that background points reflected the spatial structure of observation effort while remaining ecologically relevant to each species.

We first pooled all cleaned occurrence records within each major vertebrate group, namely amphibians, birds, mammals, and reptiles, and used these pooled records to construct a taxon-specific sampling-bias surface. For each taxon, occurrence density was smoothed using a Gaussian kernel to generate a continuous raster representing relative sampling intensity on the 10 km analysis grid (Fig. S2). This step provided a spatially explicit estimate of where records were more likely to have been collected and therefore where background sampling should be more concentrated. To avoid selecting background points from regions that were clearly outside the potential environmental domain of a species, the taxon-level bias surface was clipped for each species using a 500 km buffer around its expert-derived geographic range. This restriction limited background sampling to areas close to the known distribution and reduced the risk of inflating model performance by contrasting presences with environmentally unrealistic regions that the species is unlikely to have accessed. Species-specific background points were then drawn randomly from the clipped bias surface, with sampling probability weighted by the local bias value. In this way, areas with greater apparent sampling effort contributed proportionally more candidate background points, allowing the background sample to better mirror the spatial bias structure of the occurrence data. For most species, up to 10000 background points were generated. For species with very restricted buffered ranges, the number of sampled points was reduced so that the background set remained proportional to the available grid cells.

This taxon-specific, bias-informed background sampling procedure was used for all species included in climatic suitability modelling. By incorporating the spatial structure of survey effort into background selection, the approach reduced the influence of uneven occurrence density and improved the robustness of the resulting climatic suitability estimates.

Supplemental figures

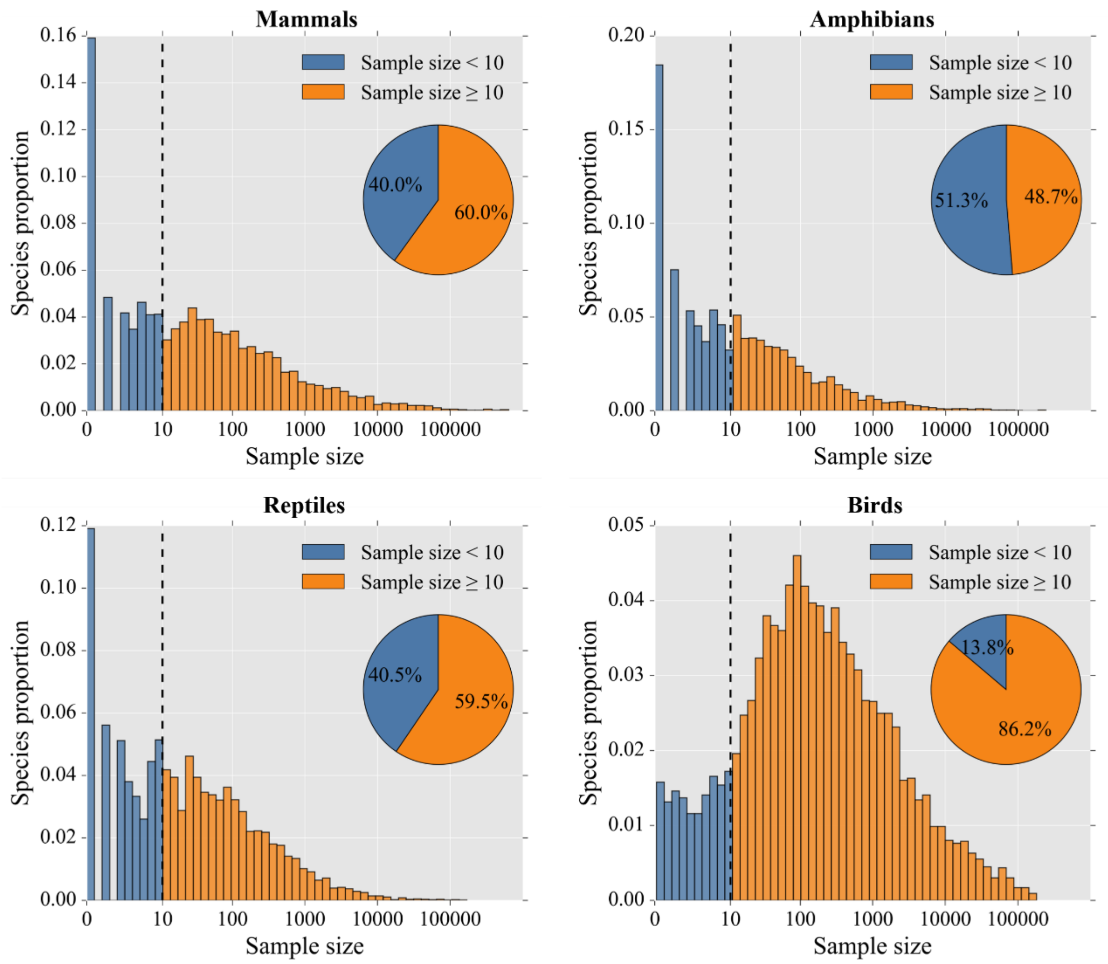

Fig. S1 Frequency distributions of effective occurrence sample sizes for four terrestrial vertebrate taxa*. Effective occurrences were defined as records located within the 500 km buffer of each species' IUCN range after removing duplicates with identical spatial coordinates and collection year. Histograms show the distribution of species by sample size, the dashed line marks the threshold of 10 samples, and the pie charts summarize the proportions of species with sample sizes < 10 and ≥ 10.*

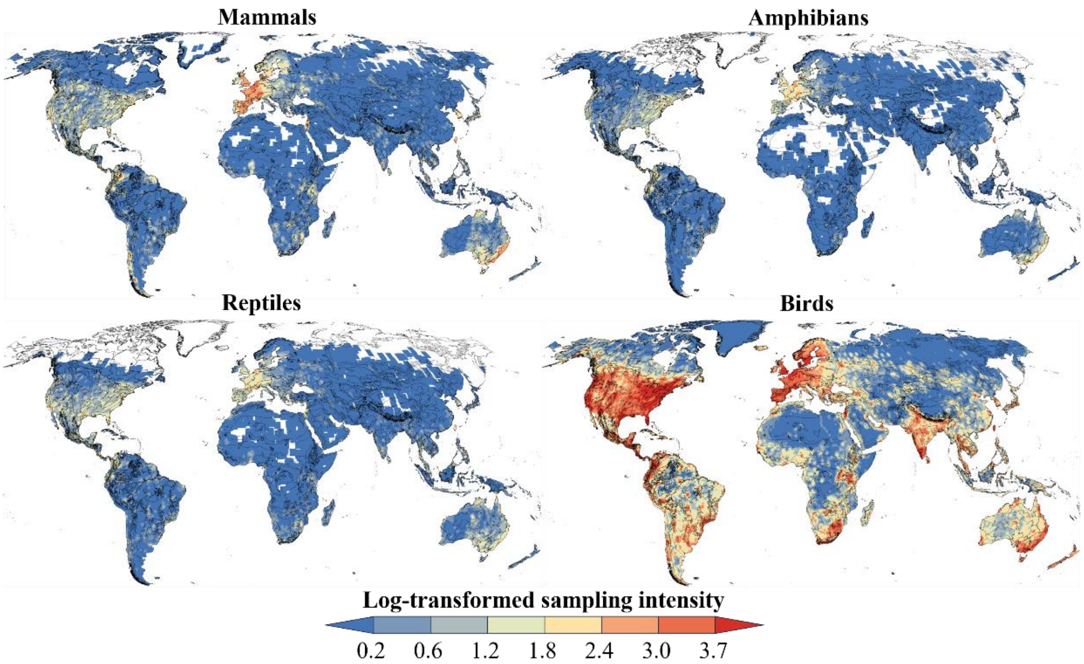

Fig. S2 Log-transformed sampling-bias surfaces for four terrestrial vertebrate taxa. *Sampling-bias surfaces were generated by rasterizing occurrence records onto the target grid, counting the number of records in each pixel, and then applying a Gaussian filter to the count raster. The smoothed values were subsequently transformed as log10(x + 1) for visualization, where x is the original Gaussian-smoothed sampling intensity.*

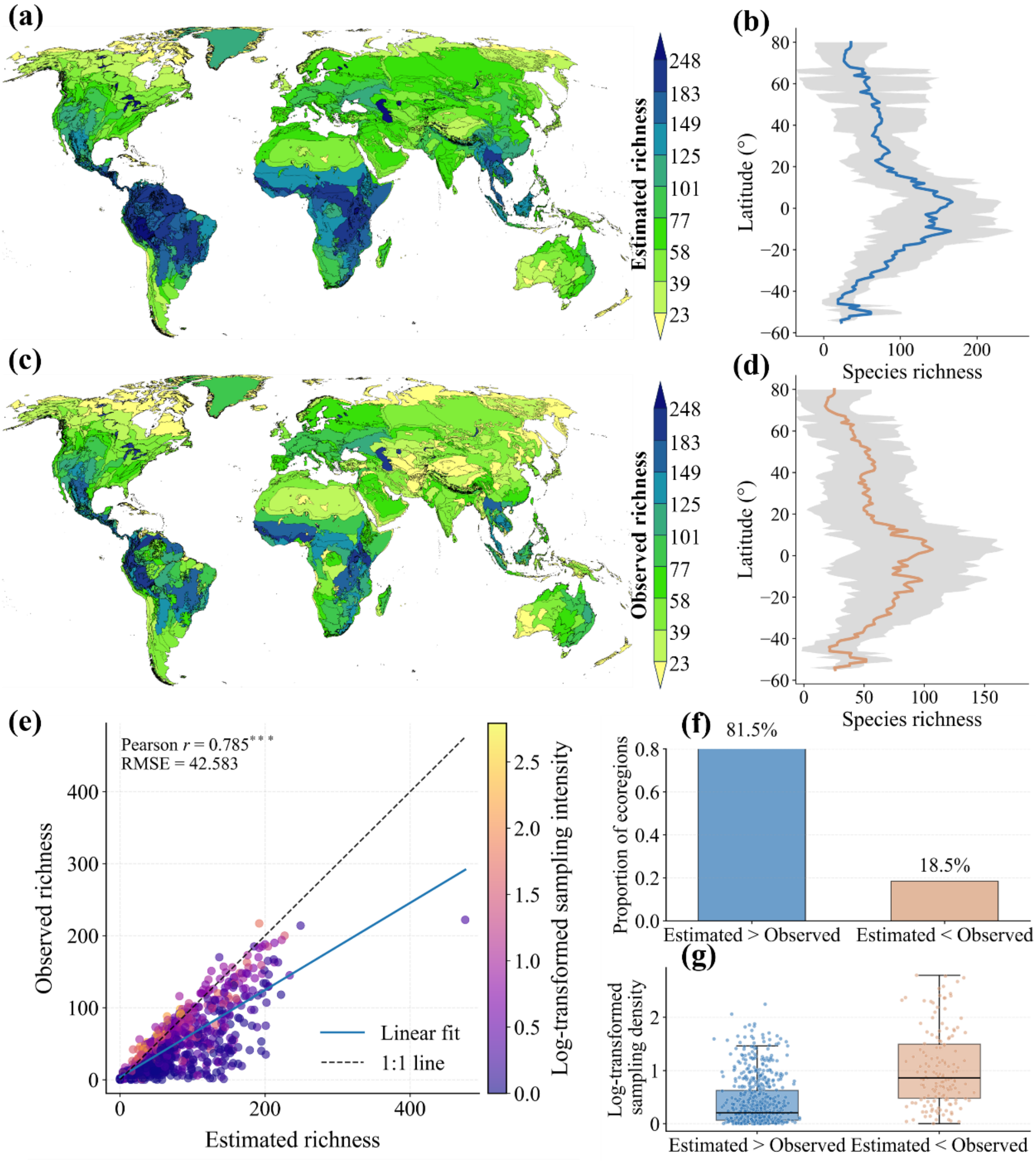

Fig. S3 Dataset-level validation of global estimated mammal species richness against occurrence-derived mammal richness across terrestrial ecoregions.

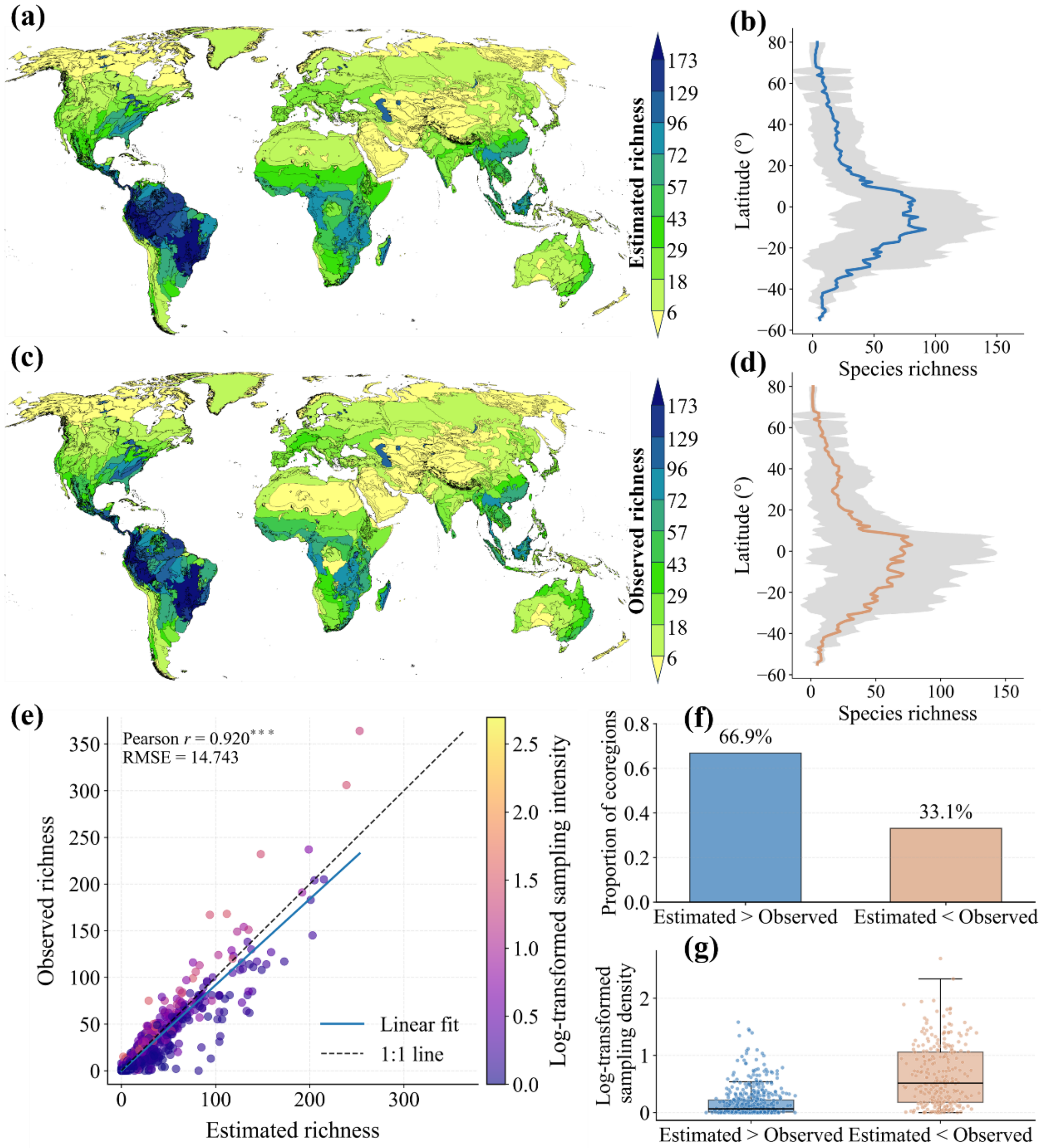

Fig. S4 Dataset-level validation of global estimated amphibian species richness against occurrence-derived amphibian richness across terrestrial ecoregions.

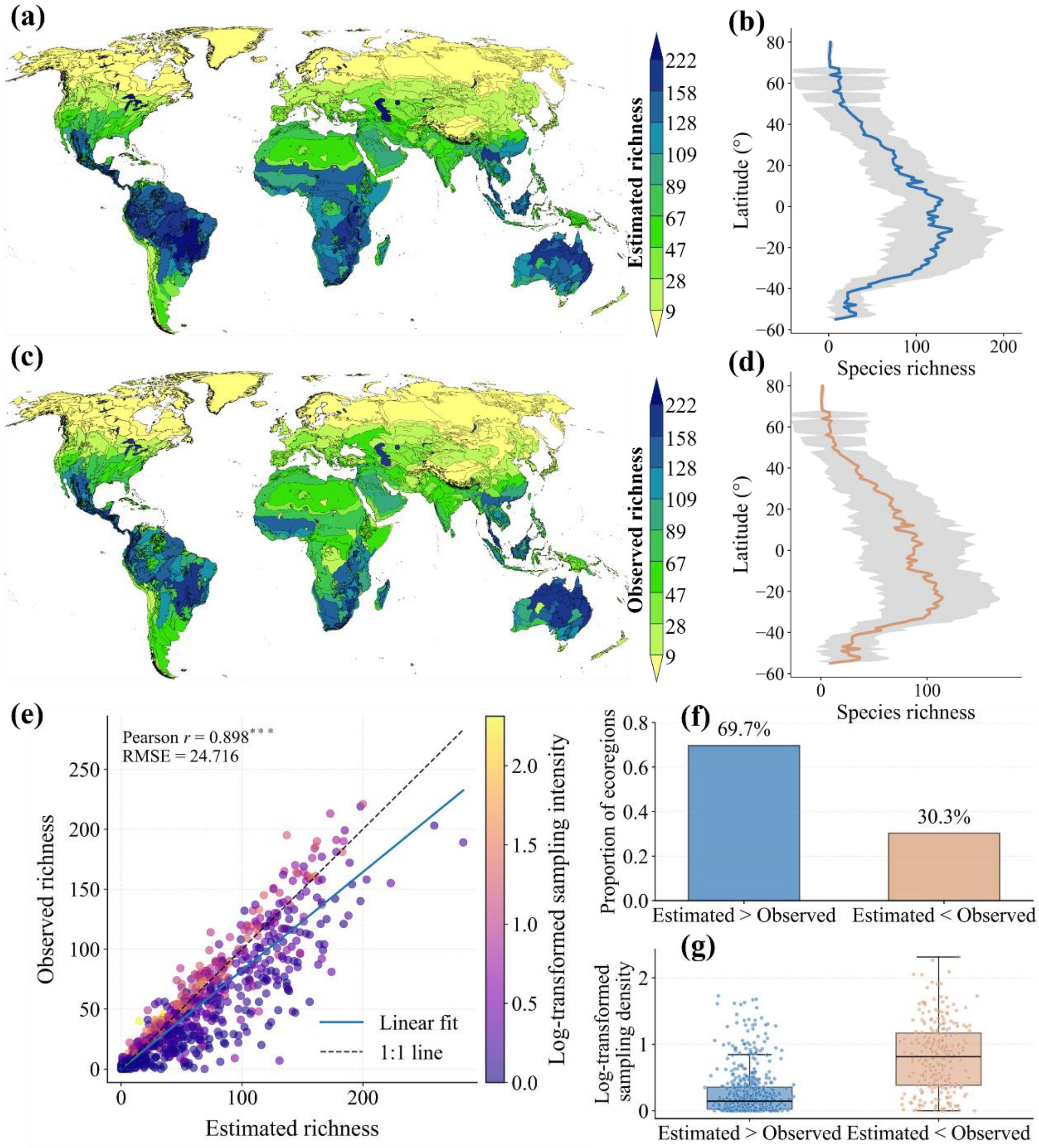

Fig. S5 Dataset-level validation of global estimated reptile species richness against occurrence-derived reptile richness across terrestrial ecoregions.

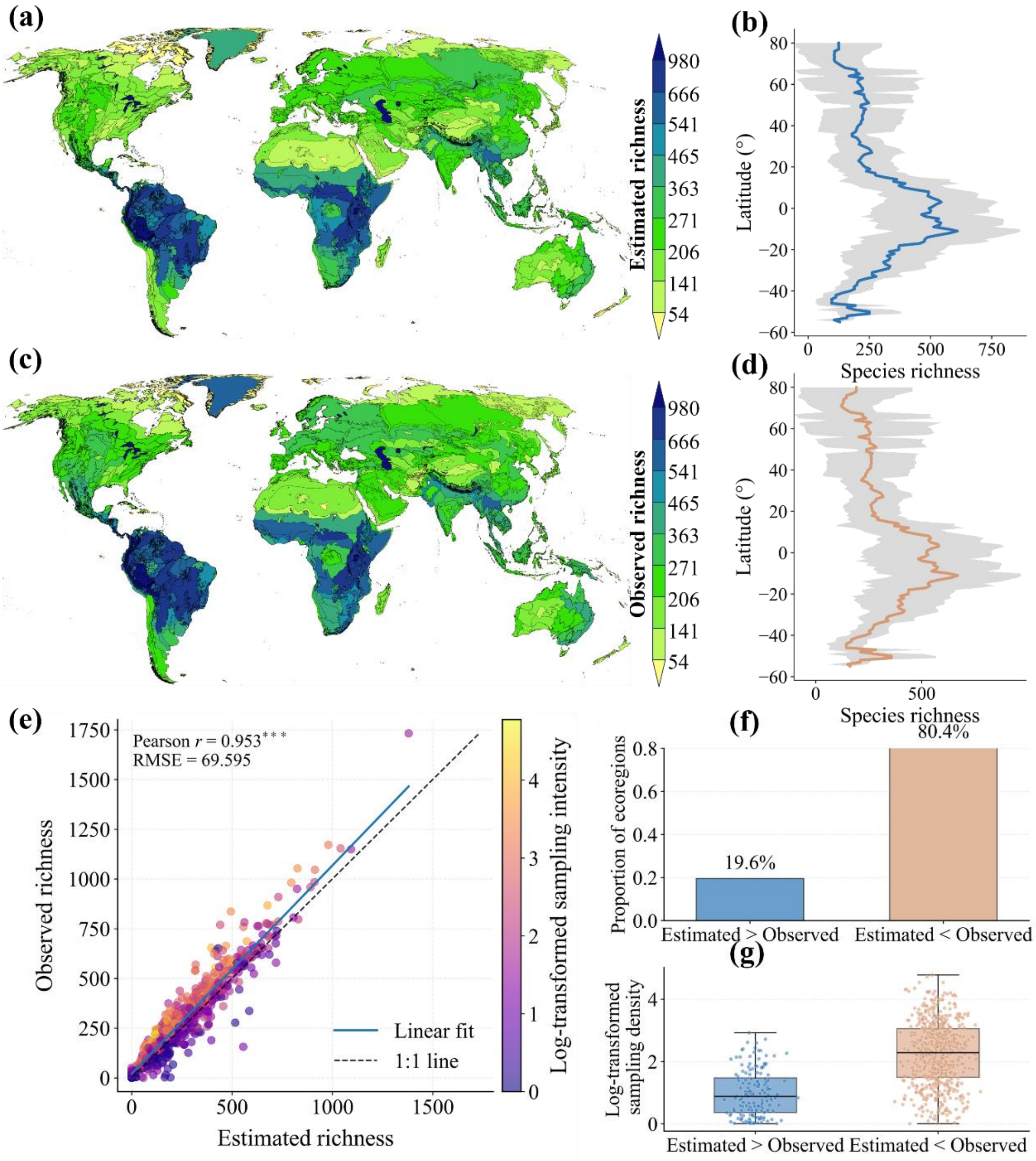

Fig. S6 Dataset-level validation of global estimated bird species richness against occurrence-derived bird richness across terrestrial ecoregions.

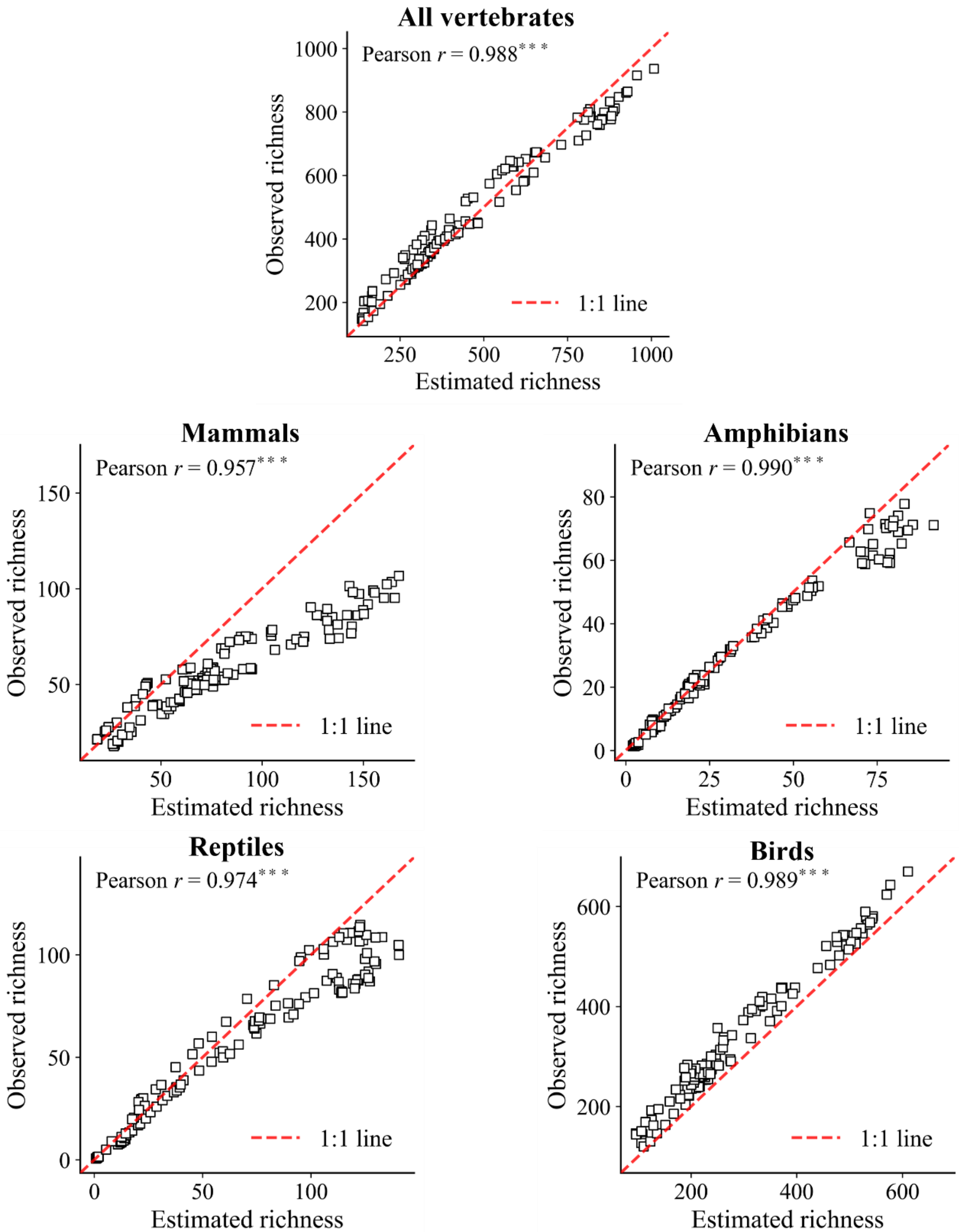

Fig. S7 Latitudinal comparison of observed and estimated species richness. *Each point represents the mean richness of all ecoregions intersecting a given 1° latitudinal band between 57° S and 80° N. This comparison was used to evaluate the similarity in spatial patterns between estimated and observed results.*

Supplemental tables

Table S1 Cross-walk between land-cover categories and IUCN habitat classes used to reconstruct temporally explicit habitat maps

Habitat category	IUCN habitat type	PFT vegetation function type	From GLC Plus LULC type
Forest	1.1. Forest – Boreal 1.2. Forest – Subarctic 1.3.Forest – Subantarctic 1.4. Forest – Temperate 1.5.Forest– Subtropical/tropical dry 1.6.Forest – Subtropical/tropical moist lowland 1.7.Forest– Subtropical/tropical mangrove vegetation above high tide level 1.8.Forest – Subtropical/tropical swamp 1.9. Forest – Subtropical/tropical moist montane 14.3 Plantation	2. Broadleaf evergreen tree, tropical 3. Broadleaf evergreen tree, temperate 4. Broadleaf deciduous tree, tropical 5. Broadleaf deciduous tree, temperate 6. Broadleaf deciduous tree, boreal 7. Needleleaf evergreen tree, temperate 8. Needleleaf evergreen tree, boreal 9. Needleleaf deciduous tree	2. Forest
non-forest natural vegetation,	2.1. Savanna - Dry 2.2. Savanna - Moist 3.1. Shrubland – Subarctic 3.2. Shrubland – Subantarctic 3.3. Shrubland – Boreal 3.4. Shrubland – Temperate 3.5. Shrubland – Subtropical/tropical dry 3.6. Shrubland – Subtropical/tropical moist 3.7. Shrubland – Subtropical/tropical high altitude 3.8. Shrubland – Mediterranean-type shrubby vegetation 4.1. Grassland – Tundra 4.2. Grassland – Subarctic 4.3. Grassland – Subantarctic 4.4. Grassland – Temperate 4.5. Grassland – Subtropical/tropical dry 4.6. Grassland – Subtropical/tropical seasonally wet/flooded	10. Broadleaf evergreen shrub, temperate 11. Broadleaf deciduous shrub, temperate 12. Broadleaf deciduous shrub, boreal 13. C3 grass, arctic 14. C3 Grass 15. C4 Grass 16. Mixed C3/C4 grass	3. Grassland 4. Shrubland 5. Tundra

	4.7. Grassland – Subtropical/tropical high altitude 5.1. Wetlands (inland) – Permanent rivers/streams/creeks (includes waterfalls) 5.2. Wetlands (inland) – Seasonal/intermittent/irregular rivers/streams/creeks 5.3. Wetlands (inland) – Shrub dominated wetlands 5.4. Wetlands (inland) – Bogs, marshes, swamps, fens, peatlands 5.5. Wetlands (inland) – Permanent freshwater lakes (over 8 ha) 5.6. Wetlands (inland) – Seasonal/intermittent freshwater lakes (over 8 ha) 5.7. Wetlands (inland) – Permanent freshwater marshes/pools (under 8 ha) 5.8. Wetlands (inland) – Seasonal/intermittent freshwater marshes/pools (under 8 ha) 5.9. Wetlands (inland) – Freshwater springs and oases 5.10. Wetlands (inland) – Tundra wetlands (inc. pools and temporary waters from snowmelt) 5.11. Wetlands (inland) – Alpine wetlands (inc. temporary waters from snowmelt) 5.12. Wetlands (inland) – Geothermal wetlands 5.13. Wetlands (inland) – Permanent inland deltas 5.14. Wetlands (inland) – Permanent saline, brackish or alkaline lakes 5.15. Wetlands (inland) – Seasonal/intermittent saline, brackish or alkaline lakes and flats 5.16. Wetlands (inland) – Permanent saline, brackish or alkaline marshes/pools 5.17. Wetlands (inland) – Seasonal/intermittent saline, brackish or alkaline marshes/pools 5.18. Wetlands (inland) – Karst and other subterranean hydrological systems (inland) 6. Rocky Areas 8.1. Desert – Hot 8.2. Desert – Temperate 8.3. Desert – Cold		
agriculture	14.1 Arable land 14.2 Pasture 14.3 Plantation 14.4 Rural gardens	18. Cropland	1. Cropland
urban	14.5 Urban	19. Urban	8. Impervious surface

Table S2 Crosswalk between the original 14 biomes and the aggregated 8 biome groups

Current biome name	**Original biome ID(s)**	**Original biome name(s)**
Tropical forests	1 2 3 14	Tropical and Subtropical Moist Broadleaf Forests; Tropical and Subtropical Dry Broadleaf Forests; Tropical and Subtropical Coniferous Forests; Mangroves
Temperate forests	4 5	Temperate Broadleaf and Mixed Forests; Temperate Conifer Forests
Boreal forests	6	Boreal Forests/Taiga
Tropical grasslands	7 9	Tropical and Subtropical Grasslands, Savannas and Shrublands; Flooded Grasslands and Savannas
Temperate grasslands	8 10	Temperate Grasslands, Savannas and Shrublands; Montane Grasslands and Shrublands
Tundra	11	Tundra
Mediterranean	12	Mediterranean Forests, Woodlands and Scrub
Deserts and Xeric Shrublands	13	Deserts and Xeric Shrublands

Supplemental references

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

Jung, M., Dahal, P.R., Butchart, S.H., Donald, P.F., De Lamo, X., Lesiv, M., Kapos, V., Rondinini, C. and Visconti, P., 2020. A global map of terrestrial habitat types. Scientific data, 7(1): 256.

Yu, L., Du, Z., Li, X., Gu, J., Li, X., Zhong, L., Duojiweise, Wu, H., Zhao, Q. and Ma, X., 2025. FROM-GLC Plus 3.0: Multimodal Land Change Mapping with SAM and Dense Surface Observations. Journal of Remote Sensing, 5: 0728.