Title: Emergent properties and robustness of species-habitat networks for global terrestrial vertebrates

Running title: species-habitat networks of terrestrial vertebrates

Abstract

Aim.

Habitat loss is the dominant cause of biodiversity decline around the world, yet the complexity and stability of terrestrial assemblages related to suitable habitats have been almost unknown on a global scale.

Location.

Global.

Time period.

Contemporary.

Major taxa studied

Mammalia, Aves, Reptilia, Amphibia.

Methods

We constructed gridded maps of species-habitat networks of terrestrial vertebrates based on global species distributions and a recently-developed habitat type dataset. Then we investigated the biogeographic patterns of emergent network structures and analyzed network robustness to habitat loss by simulating habitat removals on a global scale.

Results

We found that, compared with reptiles and amphibians, the species-habitat networks

of mammals and birds were characterized by higher habitat diversity, connectance and modularity. All four taxonomical groups have high robustness globally, but after adjusting for species and habitat diversity, we found a variation of surplus and deficiency of network structures and robustness. Temperature and precipitation contributed most to relative network robustness globally, whereas geographical and human population factors played important roles in scattered regions on all continents.

Main conclusions

Overall, we provide novel insights into the biogeographic patterns of species-habitat connections through a network approach, which can help to identify gaps for reestablishing species-habitat links to improve conservation outcomes.

Keywords: species-habitat network, network structure, habitat loss, robustness, biogeography, macroecology

Introduction

Global biodiversity is declining at unprecedented rates in human history, with several studies suggesting we are in the middle of a sixth mass extinction event (Barnosky *et al.*, 2011; Ceballos *et al.*, 2015; Cowie *et al.*, 2022). The massive loss of species threatens natural ecosystem function and stability, with widespread consequences for humans and other life on Earth (Raffaelli, 2004; Potts *et al.*, 2010). Losing habitat for a species can commit the species to extinction (Brooks *et al.*, 2002), and it is well known that habitat loss has become the dominant driver for biodiversity loss worldwide (Giam *et al.*, 2010; Pimm *et al.*, 2014; Powers & Jetz, 2019). Therefore, increasingly habitats themselves are in the center of conservation focus (Jung *et al.*, 2020; Keith *et al.*, 2022), and it is essential to know how species are connected to different types of habitats.

Species exhibit significant variation in their ranges, giving rise to broad-scale patterns in species richness along environmental gradients (Rahbek, 2005; Fine, 2015). Given that species have habitat preferences and can only persist in their suitable habitats, species' geographical ranges (such as spatial ranges from IUCN) can be refined by their preferred habitats, known as area of habitat (AOH) (Brooks *et al.*, 2019; Jung *et al.*, 2021). Habitat diversity, defined as the number of habitats in each territorial unit (Rosenzweig *et al.*, 2003; Hortal *et al.*, 2009), has been recently highlighted as a spatially explicit aspect of biodiversity, which drives species diversity across a variety of taxa. Similar to species richness, habitat diversity itself also varies with environmental conditions (Cervellini *et al.*, 2021). Although global patterns of

both species and habitat diversity have been well investigated, the interactions between them are largely unknown from a biogeographic perspective, and a biogeographic network approach or thinking might facilitate further understanding of this relationship.

Species-habitat networks have recently been investigated in studies conducted on small spatial scales where complete species assemblages were sampled across sites (Nardi *et al.*, 2019; Astudillo *et al.*, 2020; Palmeirim *et al.*, 2022). However, large-scale studies can help to reveal the spatial variation of biodiversity broad-scale ecological patterns and processes along geographic gradients (McGill, 2019). Metrics of network structure, mostly utilized in previous studies on species interactions, can also inform biodiversity assessments at broader scales (Doherty *et al.*, 2023). A classic index of network complexity is connectance, defined as the actual number of links divided by the number of possible links. It was shown the richness- and connectance-related structures were uncorrelated in space and were driven by different environmental factors in European food webs (Braga *et al.*, 2019). Other network emergent properties such as modularity also exhibited significant spatial variations globally and were mediating community resilience to disturbance (Liu *et al.*, 2021). In a similar context, species and their habitats, in terms of their occurrence and amount, are both expected to vary globally with natural conditions and human disturbance (Rahbek, 2005; Fine, 2015; Cervellini *et al.*, 2021), which inevitably leads to spatial variation of multi-level structures of species-habitat networks. For example, species-level habitat specialization can imply how species use a range of resources across

multiple habitat types, and connectance can indicate whether a species-habitat network is dominated by habitat generalists or specialists (Lami *et al.*, 2021). Ultimately a network approach could thus lead us towards a better understanding of community responses to habitat composition and changes. Currently, the broad-scale variations of species-habitat networks and the associated drivers have not been investigated, leading to a knowledge gap on how terrestrial community respond to the global threats of species habitats from climate change and human encroachment.

Based on global species distribution and a recently-developed habitat type dataset (Jung *et al.*, 2020), we constructed gridded maps of species-habitat networks of terrestrial vertebrates and investigated the biogeographic patterns of network structures and robustness on a global scale. Here, a species-habitat network refers to a bipartite network with two sets of nodes: species and their suitable habitats (Fig. 1a-e), and network robustness refers to species' capacity to withstand the removal of habitats in a network. Because of their distinct physiological systems and behavioral flexibility, the four major classes of terrestrial vertebrates have adapted to various habitats and might exhibit different degrees of habitat generalism/specialism across the globe, which might shape the organization of species-habitat links in different ways. Thus we expected taxonomical classes to differ in the average levels and spatial patterns of network structures. The robustness can be indicative of the tolerance of species pools subject to habitat disturbance, and we expect that robustness varies with environmental gradients due to the heterogeneous spatial distribution of species. Thus, we also mapped network robustness and identified the dominant environmental

103 factors affecting them on the global and regional scales.

104

105 **Methods**

106 *Construction of terrestrial species-habitat networks*

107 Similar to the habitat refinement of species distribution undertaken by Jung *et al.*
108 (2021), we formatted the connections between species and habitats as bipartite
109 networks through habitat preference (Fig. 1a). The two parties of nodes in a network
110 are species and habitats respectively.

111 The spatial range data of mammals and amphibians were retrieved from the
112 International Union for Conservation of Nature (IUCN, 2022). The spatial range data
113 of birds was requested from BirdLife International (2023), which was compiled by
114 BirdLife International and Handbook of the Birds of the World. The spatial range data
115 of reptiles was obtained from GARD 1.7 dataset (Global Assessment of Reptile
116 Distributions), which is an updated version of global reptile distributions and collated
117 from various sources and had expert validation and curation (Roll *et al.*, 2017). All
118 vertebrate spatial range maps were pre-processed following common practice by
119 selecting only those parts of a species' range where (1) it is extant, (2) it is native,
120 reintroduced or assisted colonization, and (3) it is resident or breeding distribution
121 (Jetz *et al.*, 2012; Jenkins *et al.*, 2013). We focused on terrestrial vertebrate species
122 and subsequently excluded those that live only in marine habitats. The nonbreeding
123 ranges and extinct species were also excluded.

124 We used a recently developed global habitat database with a spatially explicit

characterization of terrestrial habitat types (Jung *et al.*, 2020). The terrestrial habitat classes were delineated by intersecting data on land cover, climate and land use, and they are in the broadest sense consistent with IUCN classification, and can assist with the quantification of area of habitat estimates (Brooks *et al.*, 2019). The habitat data have two levels: level 1 and level 2 (see Table S1 for details). For example, '1. Forest' is a level-1 habitat, and the level-2 habitats include '1.1. Forest – Boreal', '1.2. Forest – Subarctic', '1.3. Forest – Subantarctic', '1.4. Forest – Temperate', etc. Because level-1 data led to very few habitat nodes (mostly one node) at 1-degree grid cells and provided little information on species-habitat links, we used the level-2 habitats as habitat nodes for constructing species-habitat networks.

The species-habitat networks were built at 1-degree grid cell which was common in previous global-scale studies, providing a compromise between omission and commission errors in the range maps (Jetz *et al.*, 2008) as well as allowing a sufficient number of habitat nodes to construct networks along environmental gradients. By matching the species spatial range maps and the habitat map, we generated the species pool and habitat composition in each grid cell. A habitat with less than 0.01% of the area was removed from each cell, which excluded those species with unrealistic areas of habitats. To test the sensitivity of our results to area thresholds, 1% and 0.0001% were also considered. Then, according to the species' preferred suitable habitat types from the IUCN Red List, the species nodes and habitat nodes were linked to form bipartite species-habitat networks for each vertebrate class in grid cells. The species or habitats that had no links were removed from the grid cells.

147 *Network structure metrics*

148 In this study, we investigated the interlinkages between species pools and the
149 occurrence of their habitats at a macroecological scale. In contrast to studies that
150 quantified species interaction networks, we used network metrics to calculate where
151 and how species were interlinked with habitats. Network size was usually the number
152 of all nodes in networks. However, species nodes and habitat nodes have distinct
153 ecological meanings, and adding them together made little ecological sense. We
154 therefore used habitat diversity (number of habitat nodes) and species diversity
155 (number of species nodes) as two components of network size for a species-habitat
156 network. Preliminary analysis indicated habitat diversities (H) were moderately
157 correlated with species diversities (S) (Pearson correlation = 0.6 for mammals, 0.69
158 for birds, 0.6 for reptiles, 0.53 for amphibians, all $P < 0.001$) and have been already
159 mapped in previous studies (www.iucnredlist.org) (Fig. S2). Thus, we mapped the
160 relative habitat diversity in each cell using the residuals from the regression, habitat
161 diversity ~ rarity-weighted species diversity. The rarity-weighted species diversity
162 combines endemism and species diversity (Fig. S2), calculated by summing the rarity
163 values of all species in a cell. For each species, the rarity value is the proportion of the
164 species' range contained within that cell. The maps of relative habitat diversity
165 showed the areas where one would expect fewer or more habitats according to the
166 number of geographically rare (small-ranged) species.

167 We then measured network connectance and modularity that were commonly
168 used in literature and showed relationships with network robustness (Dunne *et al.*,

169 2002; Landi *et al.*, 2018; Liu *et al.*, 2021; Palmeirim *et al.*, 2022). Connectance was
170 measured as the realized proportion of species-habitat links (L) given all species
171 occurred in all habitats, C = L/ (S×H). A network with high connectance was
172 dominated by habitat generalists, and one with low connectance was dominated by
173 habitat specialists. In realistic networks like food webs and bipartite mutualistic
174 networks (Yodzis, 1980; Thébault & Fontaine, 2010), connectance is often negatively
175 correlated to network size, thus comparison of this metric needs to take the
176 dependence on network size into account (Bengtsson, 1994). Our preliminary analysis
177 found connectance was also strongly correlated to species and habitat diversity, so we
178 performed log-regression of the connectance against species diversity, habitat
179 diversity and their interaction terms, and used the residuals to represent the relative
180 connectance, which represents the degree of the higher or lower connectance than
181 expected for network size. Modularity tests the tendency of whether a network can be
182 divided into subgroups in which the nodes strongly interact among themselves
183 compared to other nodes within the network (Newman, 2006) (Fig. 1f). It has been
184 shown to buffer the spread of a perturbation across the entire network (Gilarranz *et*
185 *al.*, 2017), and thus could potentially buffer species extinctions from habitat loss in
186 species-habitat networks. We performed modularity analysis by using
187 *computeModules* from the *bipartite* package in R software (Dormann *et al.*, 2008;
188 Beckett, 2016). Because modularity is not only related to network size and links, but
189 also related to nodes' degree sequences, we standardized modularity relative to null
190 models by calculating the relative modularity (Sebastián-González *et al.*, 2015; Liu *et*

191 *al.*, 2021), (observed modularity - μ_{null})/ μ_{null}, where the μ_{null} represents the mean of
192 modularity values from 100 null models. We chose to use the "quasiswap" null model
193 that generated random networks while preserving the degrees for both species and
194 habitat nodes, from the "vegan" package in R software. This null model is also called
195 fixed-fixed algorithm for randomizing matrices and often used in bipartite network
196 studies investigating modular patterns (Olesen *et al.*, 2007; Thébault, 2013). To test
197 the sensitivity of modularity to null models, we also applied two other null models:
198 the "r1" null model algorithm preserves row totals (species number in each habitat)
199 and uses column totals as probabilities of selecting species (species's habitat
200 generalism); the "c1" null model algorithm preserves column totals (habitat number
201 for each species) and uses row totals as probabilities of selecting habitat. The other
202 two null models can potentially create degenerate matrices, that is, matrices in which
203 one or more rows or columns are empty.
204 *Robustness to habitat loss*
205 Robustness is a widely used topological stability measure indicating the resilience of a
206 network to node loss (Landi *et al.*, 2018). The network robustness in this study was
207 interpreted as how the proportion of remaining species changed after the simulated
208 removal of habitats from the species-habitat network (Fig. 1g). We assumed that a
209 species went extinct in a given grid cell if all its suitable habitats were lost and a
210 losing habitat type could not be recovered or replaced by other suitable habitats
211 during the simulation. To assess this, we tested both random and sequential removal
212 scenarios. The random scenario was tested as a baseline, in which the habitat was

213 removed randomly each time until no habitat existed. The sequential scenarios were
214 applied by removing habitats ordered by the total area of each habitat: one from the
215 lowest to largest ones (small-to-large) and the other from the largest to lowest ones
216 (large-to-small). Because large-sized habitats are thought to be more important for
217 species persistence than small ones, they are usually better maintained and less
218 disturbed (Tulloch *et al.*, 2016). To put it another way, smaller habitats are easier to
219 lose than larger ones. Therefore, we assumed the low-to-large scenario was more
220 realistic than the large-to-small and random scenarios, and focused on the robustness
221 of this scenario in the following analysis. The raw robustness is calculated as the area
222 below the species extinction curve (Fig. 1g).
223 Because the robustness was highly correlated with species and habitat diversity,
224 its spatial pattern was redundant with that of diversities. We defined the relative
225 robustness using the residuals of regressions of robustness against species diversity,
226 habitat diversity and their interaction terms. The relative robustness revealed the
227 surpluses and deficits compared to the expected robustness of species pools with
228 given species and habitat diversity, indicating whether a species pool would face
229 lower- or higher-than-expected tolerance to habitat loss.
230 *Environmental variables*
231 We used climatic, geographical and human population density to test the spatial
232 change of relative robustness on environmental gradients. The 10-minute maps of
233 climate variables including annual mean temperature (in degrees Celsius), annual
234 precipitation (in millimeters), temperature seasonality, and precipitation seasonality

were downloaded from worldclim.org. The 30 arc-second elevation data (in meters) was downloaded from GLOBE database (Hastings *et al.*, 1999). The topography was measured by the terrain ruggedness index, which is the mean of the absolute differences in elevation, downloaded from Amatulli *et al.* (2018). The 1-degree human population density (persons per square kilometers) data for the years 2000, 2005, 2010, 2015, and 2020 were downloaded from SEDAC database (GPWv4, 2016). The mean value of human densities over 5 timeslices in each cell was calculated to indicate average human disturbance over the last two decades. All the maps of environmental variables were downscaled to the spatial resolution of species-habitat networks.

Statistical analysis

We have three parts of statistical tests. First, we used Kruskal-Wallis rank sum test to test the difference of average network structures or robustness among four taxonomic classes, followed by Dunn's tests for multiple comparisons. Second, Spearman correlation analysis was performed to test the correlation of each network structure among classes and the correlation among network structures in each class.

Third, random forest (RF) and linear models (LM) were applied to test the associations of relative small-to-large robustness with environmental variables for each class of species on the global and regional scales. RF is a commonly used machine learning model, which has advantages such as nonlinear effect, high prediction accuracy, and handling high-dimension large datasets in ecology (Prasad *et al.*, 2006; Cutler *et al.*, 2007). Especially, the variable importance measure in RF can

be used to subjectively identify ecologically important variables for interpretation. We used the permutation importance that measures the difference in prediction accuracy before and after permuting values of the variable, averaged over all trees (Nicodemus *et al.*, 2010). On the global scale, all available gridded cells were included in the RF and LM. In both RF and LM, relative robustness was the response variable, and environmental variables were independent variables. Elevation, ruggedness, precipitation seasonality and human population density were log-transformed to reduce the skewness, and temperature seasonality was excluded due to its high correlation with annual temperature. We also checked that multicollinearity was low by variance inflation factor (VIF) (all VIF < 2.1) in linear models (Table S2). The fitting performance of RF and LM was evaluated by the coefficient of determination (R^2). Due to the poor fitting of linear models and nonlinear associations detected in multiple variables by RF models, we focused on the results of RF models. We also tested the associations between the unadjusted robustness and environmental variables using RF models with species and habitat diversities as covariates.

We applied the local version of RF models, also called the geographically weighted random forest model (GWRF), to analyze the spatial variation of dominant variables determining the relative robustness on the regional scale. The GWRF allows for the investigation of the existence of spatial non-stationarity in the relationship between dependent and independent variables. For this algorithm, we chose 100 nearest neighbor cells (bandwidth = 100) and 2000 trees. For global and regional RF models, we showed partial dependence plots that visualized the relationship between

the response variable and environmental variables. The above analyses were performed using the "*ranger*" and "*SpatialML*" packages in R 4.1.2 software (Wright & Ziegler, 2017; Georganos *et al.*, 2021; R-Core-Team, 2022).

Results

In total, we mapped 5162 mammals occurring across 49 habitats, 9807 birds in 50 habitats, 7944 reptiles in 44 habitats, and 6367 amphibians in 44 habitats around the world (Fig. 1 and Table S1). Forested habitats (such as subtropical/tropical, moist/dry, lowland/montane) were the habitats with most species, followed by artificial terrestrial areas (Fig. S1, including rural gardens, plantations, pasturelands, etc.) and inland wetlands. Overall, Kruskal-Wallis tests suggest all the tested network structures and robustness had significant differences among four classes ($P < 0.001$), and the effect sizes varied with metrics (Fig. 2). On average, reptiles and amphibians had lower habitat diversity than mammals and birds (Fig. 2a). On average, the connectance was lowest for birds, and highest for amphibians (Fig. 2b). The average modularity values of birds and mammals were higher than that of reptiles and amphibians by three null models. (Fig. 2c and Fig. S3). Finally, there were significant differences among the three scenarios for robustness, ordered as small-to-large > random > large-to-small scenarios (Fig. 2e).

Spatial patterns of network structures

The network metrics of species-habitat networks all showed strong spatial variation globally (Fig. 3, Fig. S4-S6) and were positively correlated among four terrestrial classes in space (Fig. S7). The areas with high relative habitat diversity were mostly located in North America (mainly for amphibians and reptiles), Europe and North Asia, while those with low relative habitat diversity were mainly in Australia, Sahara, Tibetan Plateau and North Canada (Fig. 3a, Fig. S4), suggesting the habitat types were fewer than expected for the geographically rare species in these regions.

Connectance was negatively correlated to species and habitat diversity in all four terrestrial classes (Fig. S8). The relative connectance was mostly close to zero globally, suggesting most networks had expected connectance levels as species and habitat diversity predicted. The networks with higher relative connectance were mainly located in Europe and Africa, while those with lower relative connectance were mainly located in North and Southeast Asia (Fig. 3b, Fig. S5).

The networks with high relative modularity from the "quasiswap" null model were mainly located in the USA, South Asia, East Africa and mountainous areas in South America for four taxonomical classes, but the reptiles and amphibians showed lower modularity than mammals and birds in Europe (Fig. 3c, Fig. S6). The modularities from the other two null models were positively correlated with that from "quasiswap" model in spatial patterns (Table S3), but the results from the "r1" null models were slightly different, with widespread high modularities in North America and Eurasia for mammals and birds (Fig. S9-S10).

Spatial patterns of robustness

The unadjusted robustness from three scenarios was positively correlated in space (Fig. S8 and S11). The small-to-large (unadjusted) robustness was generally high in most regions (Fig. 4a), and a few species-habitat networks with low robustness were mainly located in low-diversity areas such as Tibetan Plateau, Sahara, and central Australia. After adjustment by species diversity and habitat diversity, the relative robustness showed greater spatial variations worldwide (Fig. 4b). The spatial variations of relative robustness were weakly and positively correlated among four classes (Fig. S7e). For mammals and birds, the species-habitat networks with low relative robustness were mainly located in the Americas and middle-latitude Asia; Australia has also low relative robustness for reptiles and amphibians (Fig. 4b).

The R^2 of Global RF models of relative robustness were around 50% in four taxonomic groups (R^2 = 52% for Mammalia, 49% for Aves, 52% for Reptilia, and 45% for Amphibia). Annual temperature and precipitation were generally more important than other variables in affecting relative robustness (Fig. 4c). Annual temperature showed U-shaped effects for mammals, birds and reptiles, but a complex wiggled effect for amphibians; annual precipitation showed nonlinear positive effects except for reptiles; precipitation seasonality showed negative effects in mammals, birds and for reptiles, but a dome-shaped effect for amphibians (Fig. 5a-c). For geographical variables, elevation and ruggedness both showed dome-shaped effects for mammals, reptiles and amphibians (Fig. 5d, e). Human population density mainly showed dome-shaped effects for all four groups (Fig. 5f). The RF models for the unadjusted robustness showed similar results (Fig. S12). Linear models suggested

annual temperature showed negative effects, but annual precipitation showed positive effects in all four groups; precipitation seasonality showed negative effects in mammals and birds, positive effect in amphibians, and no effect in reptiles; elevation and ruggedness showed negative or insignificant effects in all four groups; human population density showed negative effects in mammals, birds and reptiles, but a positive effect in amphibians (Table S4).

The regional-scale RFs showed different environmental variables were dominant in different regions (Fig. 6). Temperature and precipitation were dominant in most regions around the world, whereas elevation, ruggedness and human population were dominant in scattered regions on all the continents (Fig. 6, Fig. S13). Along the latitudinal gradient, annual temperature became increasingly important from the Southern to Northern Hemisphere, and annual precipitation and ruggedness showed similar trends before the drops from middle to high northern latitudes; human population was most important at low northern latitudes for mammals and birds and at middle northern latitudes for reptiles (Fig. 6, Fig. S14).

Discussion

Biodiversity, in all its dimensions and facets, is tied to habitats, but patterns of interaction network complexity and topology between species and habitats have never been investigated on a global scale. For the first time, we have revealed the global spatial patterns of emergent network properties and robustness to habitat loss in terrestrial vertebrates by applying the species-habitat network approach. The average

levels and spatial patterns of network properties suggest the four classes of terrestrial vertebrates differed in the degrees of habitat specialism and composition. Robustness analysis further showed that the regional species pools varied in the tolerance to habitat loss across the world. Temperature and precipitation were the most important environmental factors affecting robustness globally, becoming increasingly important from the Southern Hemisphere to mid-high northern latitudes; however, human population and geography were also crucial in scattered regions worldwide. Perspectively, these results could help to identify vulnerable regions subject to habitat disturbance and loss.

Mammals and birds outperform reptiles and amphibians in thermoregulation and dispersal ability (Hickman *et al.*, 2019), which could explain their difference in suitable habitat diversity of species-habitat networks. This could also lead to higher connectance levels of reptiles and amphibians because connectance is usually negatively correlated to network size in ecological networks (Thébault & Fontaine, 2010). The relative habitat diversity showed most regions had compatible or even higher habitat diversity for rare species and the four taxa have similar spatial patterns except for some regions in America. The geographically rare species in the regions with lower-than-expected habitat diversity may be prone to extinction, such as South America (for birds), Sahara, and central Australia. Besides, how species connect habitats is also important. Connectance and modularity have been frequently investigated in species interaction networks because of their relationships with community stability and perturbation transportation (Delmas *et al.*, 2019; Mariani *et*

388 *al.*, 2019). In the context of species-habitat networks, these network topologies can
389 reveal additional information on the species pool's sensitivity to habitat change. The
390 species in the regions with low relative connectance tend to be habitat specialists. It is
391 also expected that the high modularity of species-habitat networks may lower their
392 risk of systematic collapse due to the buffering capacity of modular structures to
393 perturbations (Dormann *et al.*, 2017; Gilarranz *et al.*, 2017). Therefore, the regions
394 with lower connectance and modularity levels may be more vulnerable to habitat
395 disturbance.

396 Habitat destruction and degradation have become the dominant factors for
397 biodiversity loss. Through species-habitat networks, the simulation of sequential
398 removal of habitats can evaluate the overall risks of species pools under future habitat
399 loss. We found the robustness levels from the small-to-large scenario (small habitats
400 losing first) were higher than the random and large-to-small scenarios. This might not
401 be explained by the classic species-area relationship describes that larger areas host
402 more diverse species because we assumed a suitable habitat over an area threshold
403 could hold all the species that were expected to occur in that grid cell. Instead, our
404 results reflect that small-area habitat types occurred to be inhibited by fewer species
405 than large-area ones (Fig. S15). Furthermore, it is not surprising or new that the
406 unadjusted robustness was positively correlated with species and habitat diversity.
407 However, what is more informative is thus whether a species pool has expected
408 robustness for its species and habitat diversity. Compared with the generally high
409 levels of original small-to-large robustness, our maps of relative robustness showed

410 many species pools had lower tolerance levels to habitat loss (Fig. 4b), which can
411 guide the conservation of terrestrial habitats for vertebrates at regional scales.
412 It has been suggested that various environmental conditions, including climate,
413 geography and anthropogenic pressure, affected both species and habitat diversity at
414 large spatial scales (Rahbek, 2005; McCain, 2009; Fine, 2015; Cervellini *et al.*, 2021).
415 In this study, we found that environmental conditions could also affect the robustness
416 of species-habitat networks to habitat loss after the adjustment of species and habitat
417 diversity. The results from random forest models suggest nonlinear associations
418 between relative robustness and environmental factors, while linear models only
419 captured parts of those associations with poor fitting performance (all $R^2 < 0.06$, Table
420 S4). The associations between climate and robustness were likely caused by climate-
421 driven habitat specialization/generalization on the community level, namely, habitat
422 specialization leads to low robustness to habitat loss, and vice versa. However, two
423 contrasting hypotheses were proposed to underlie this relationship. On one hand,
424 extreme climate (high climatic seasonality, low/high rainfall and temperature) might
425 favor habitat specialists that adapted specifically to stressing environments
426 (Bastianelli et al., 2017; Rivas-Salvador et al., 2019); on the other hand, mild and
427 stable environment with high resource availability allow species subdivide the habitat
428 finely and promote coexistence of more specialized species (MacArthur *et al.*, 1966;
429 Karr, 1971; Seoane *et al.*, 2017; Rivas-Salvador *et al.*, 2019). For mammals, birds,
430 and reptiles, the effects of precipitation tend to coincide with the extreme-climate
431 hypothesis because high seasonality and low annual precipitation lead to low

robustness, while the U-shaped effect of annual temperature in amphibians likely supports the mild-climate hypothesis.

Temperature and precipitation were dominant in affecting the relative robustness of all four classes in most regions of different continents, suggesting climate predominantly shaped the connections between the species and habitats regardless of their species or habitat diversity. However, this does not mean geographical and anthropogenic factors are negligible, on the contrary, there were a considerable number of regions where they played more important roles than climate (Fig. 5). Especially, although human population density was a subordinate factor in affecting robustness, we identified a few areas that were mostly affected by human population, such as north-central Africa for mammals and reptiles, and scattered areas in Central and South America for all four groups. The dome-shaped effects of human population imply that moderate human disturbance might increase species' tolerance to habitat loss, probably providing suitable habitats for some human-adapted animals (Benton *et al.*, 2003; Fahrig *et al.*, 2011), but strong human disturbance might cause massive defaunation by destructing suitable habitats (Fischer & Lindenmayer, 2007; Buchadas *et al.*, 2023). Historically speaking, human disturbances have driven the defaunation of terrestrial vertebrates including range contractions, population declines and extinctions at local and global scales (Dirzo *et al.*, 2014), which reduced species diversity, altered regional faunal communities and might even transform habitat types through their interactions with vegetations (such as herbivory and seed predation). This might have an impact on the spatial patterns of network structures and

454 robustness, for example, some severely defaunated areas like East China and Atlantic
455 forest of South America showed low relative connectance and robustness in mammals
456 and birds (Galetti *et al.*, 2021; Ferreiro-Arias *et al.*, 2024).

457 As an attempt to analyze species-habitat interactions on a global scale, there are
458 two issues worth further investigation in our results. First, the threshold of habitat area
459 would affect the number of habitats in a network, which might affect network
460 structure and robustness. Since the exact habitat area to support species existence is
461 largely unknown for most species, we performed analyses with two other area
462 thresholds (1% and 0.0001%). The network structures and robustness using these two
463 area thresholds were highly correlated with those using the 0.01% threshold (Fig.
464 S16-S22), except that the correlation of relative modularity between 1% and 0.01%
465 was moderate (0.55-0.65) (Fig. S20), suggesting modularity was more sensitive to
466 habitat diversity than other metrics. In general, the area thresholds of habitats did not
467 affect our main results qualitatively (Fig. S23-S32). Second, the nonlinear effects
468 from global RF models were model-fitting relationships that reflected major trends of
469 relative robustness along the global ranges of environmental gradients and inevitably
470 overlooked details at regional scales. The regional RF models suggested the effects of
471 environmental variables on relative robustness were varying, being linear or nonlinear
472 within different ranges (Fig. S33-36). For example, the sharp increase of relative
473 robustness with temperature should not be interpreted as a boundless increase over
474 30 °C (Fig. 5a), because the partial dependence plots from regional RF models
475 showed flat or slightly decreasing trends within high-temperature ranges (Fig. S33-

36a). This suggests the effects of environmental variables on robustness were scale-dependent and we should be cautious to interpret them when switching between global and regional scales.

In summary, through an extension of the recent species-habitat network approach from local to global scales, we have demonstrated network metrics could uncover novel geographical patterns for connections between species and habitats. The network stability measures such as robustness can be used for studying how species pools will respond to habitat change at different environmental conditions. In future studies, realistic scenarios of habitat loss and incorporation of habitat recovery or transformation (naturally and by human intervention) will facilitate the prediction of dynamics of diversity and community stability under global climate and land use changes. We believe the network approach can add to the current paradigm of biogeographic investigations on biodiversity, and help to prioritize efforts and gaps for halting or reversing current rates of biodiversity loss.

References

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J. & Jetz, W. (2018) A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific Data*, **5**, 180040.

Astudillo, P.X., Grass, I., Siddons, D.C., Schabo, D.G. & Farwig, N. (2020) Centrality in species-habitat networks reveals the importance of habitat quality for high-Andean birds in Polylepis woodlands. *Ardeola*, **67**, 307-324.

498 Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., . . .
499 Ferrer, E.A. (2011) Has the Earth's sixth mass extinction already arrived?
500 *Nature*, **471**, 51-57.
501 Bastianelli, G., Tavecchia, G., Meléndez, L., Seoane, J., Obeso, J. & Laiolo, P. (2017)
502 Surviving at high elevations: an inter-and intra-specific analysis in a mountain
503 bird community. *Oecologia*, **184**, 293-303.
504 Beckett, S.J. (2016) Improved community detection in weighted bipartite networks.
505 *Royal Society Open Science*, **3**, 140536.
506 Bengtsson, J. (1994) Confounding Variables and Independent Observations in
507 Comparative Analyses of Food Webs. *Ecology*, **75**, 1282-1288.
508 Benton, T.G., Vickery, J.A. & Wilson, J.D. (2003) Farmland biodiversity: is habitat
509 heterogeneity the key? *Trends in Ecology & Evolution*, **18**, 182-188.
510 BirdLife International (2023) *Bird species distribution maps of the world version*.
511 Available at: http://datazone.birdlife.org/species/requestdis (accessed 2023).
512 Braga, J., Pollock, L.J., Barros, C., Galiana, N., Montoya, J.M., Gravel, D., . . . Thuiller,
513 W. (2019) Spatial analyses of multi-trophic terrestrial vertebrate assemblages in
514 Europe. *Global Ecology and Biogeography*, **28**, 1636-1648.
515 Brooks, T.M., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Rylands, A.B.,
516 Konstant, W.R., . . . Hilton-Taylor, C. (2002) Habitat loss and extinction in the
517 hotspots of biodiversity. *Conservation Biology*, **16**, 909-923.
518 Brooks, T.M., Pimm, S.L., Akçakaya, H.R., Buchanan, G.M., Butchart, S.H.M., Foden,
519 W., . . . Rondinini, C. (2019) Measuring Terrestrial Area of Habitat (AOH) and

520 Its Utility for the IUCN Red List. *Trends in Ecology & Evolution*, **34**, 977-986.
521 Buchadas, A., Jung, M., Bustamante, M., Fernández-Llamazares, Á., Garnett, S.T.,
522 Nanni, A.S., . . . Kuemmerle, T. (2023) Tropical dry woodland loss occurs
523 disproportionately in areas of highest conservation value. *Global Change*
524 *Biology*, **29**, 4880-4897.
525 Ceballos, G., Ehrlich, P.R., Barnosky, A.D., García, A., Pringle, R.M. & Palmer, T.M.
526 (2015) Accelerated modern human-induced species losses: Entering the sixth
527 mass extinction. *Science Advances*, **1**, e1400253.
528 Cervellini, M., Di Musciano, M., Zannini, P., Fattorini, S., Jiménez-Alfaro, B., Agrillo,
529 E., . . . Chiarucci, A. (2021) Diversity of European habitat types is correlated
530 with geography more than climate and human pressure. *Ecology and Evolution*,
531 **11**, 18111-18124.
532 Cowie, R.H., Bouchet, P. & Fontaine, B. (2022) The Sixth Mass Extinction: fact, fiction
533 or speculation? *Biological Reviews*, **97**, 640-663.
534 Cutler, D.R., Edwards Jr., T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. & Lawler,
535 J.J. (2007) Random forests for classification in ecology. *Ecology*, **88**, 2783-2792.
536 Delmas, E., Besson, M., Brice, M.H., Burkle, L.A., Dalla Riva, G.V., Fortin, M.J., . . .
537 Newman, E.A. (2019) Analysing ecological networks of species interactions.
538 *Biological Reviews*, **94**, 16-36.
539 Dirzo, R., Young, H.S., Galetti, M., Ceballos, G., Isaac, N.J.B. & Collen, B. (2014)
540 Defaunation in the Anthropocene. *Science*, **345**, 401-406.
541 Doherty, S., Saltré, F., Llewelyn, J., Strona, G., Williams, S.E. & Bradshaw, C.J.A.

(2023) Estimating co-extinction threats in terrestrial ecosystems. *Global Change Biology*, **29**, 5122-5138.

Dormann, C.F., Gruber, B. & Fründ, J. (2008) Introducing the bipartite package: analysing ecological networks. *R news*, **1**, 8-11.

Dormann, C.F., Frϋnd, J. & Schaefer, H.M. (2017) Identifying causes of patterns in ecological networks: opportunities and limitations. *Annual Review of Ecology, Evolution, and Systematics*, **48**, 559-584.

Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002) Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, **5**, 558-567.

Fahrig, L., Baudry, J., Brotons, L., Burel, F.G., Crist, T.O., Fuller, R.J., . . . Martin, J.L. (2011) Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. *Ecology Letters*, **14**, 101-112.

Ferreiro-Arias, I., Santini, L., Sagar, H.S.S.C., Richard-Hansen, C., Guilbert, E., Forget, P.-M., . . . Benítez-López, A. (2024) Drivers and spatial patterns of avian defaunation in tropical forests. *Diversity and Distributions*, **n/a**, e13855.

Fine, P.V.A. (2015) Ecological and evolutionary drivers of geographic variation in species diversity. *Annual Review of Ecology, Evolution, and Systematics*, **46**, 369-392.

Fischer, J. & Lindenmayer, D.B. (2007) Landscape modification and habitat fragmentation: a synthesis. *Global ecology and biogeography*, **16**, 265-280.

Galetti, M., Gonçalves, F., Villar, N., Zipparro, V.B., Paz, C., Mendes, C., . . .

564 Bovendorp, R.S. (2021) Causes and consequences of large-scale defaunation in
565 the Atlantic Forest. *The Atlantic Forest: History, Biodiversity, Threats and*
566 *Opportunities of the Mega-diverse Forest* (ed. by M.C.M. Marques and C.E.V.
567 Grelle), pp. 297-324. Springer International Publishing, Cham.
568 Georganos, S., Grippa, T., Niang Gadiaga, A., Linard, C., Lennert, M., Vanhuysse, S., . . .
569 Kalogirou, S. (2021) Geographical random forests: a spatial extension of the
570 random forest algorithm to address spatial heterogeneity in remote sensing and
571 population modelling. *Geocarto International*, **36**, 121-136.
572 Giam, X., Bradshaw, C.J.A., Tan, H.T.W. & Sodhi, N.S. (2010) Future habitat loss and
573 the conservation of plant biodiversity. *Biological Conservation*, **143**, 1594-1602.
574 Gilarranz, L.J., Rayfield, B., Liñán-Cembrano, G., Bascompte, J. & Gonzalez, A. (2017)
575 Effects of network modularity on the spread of perturbation impact in
576 experimental metapopulations. *Science*, **357**, 199-201.
577 GPWv4 (2016) Gridded Population of the World, Version 4 (GPWv4): Administrative
578 Unit Center Points with Population Estimates. Center for International Earth
579 Science Information Network - CIESIN - Columbia University. Gridded
580 Population of the World, Version 4 (GPWv4): Administrative Unit Center Points
581 with Population Estimates. Palisades, NY: NASA Socioeconomic Data and
582 Applications Center (SEDAC).http://dx.doi.org/10.7927/H4F47M2C.
583 Accessed 1 Sep. 2023. In:
584 Hastings, D.A., Dunbar, P.K., Elphingstone, G.M., Bootz, M., Murakami, H.,
585 Maruyama, H., . . . MacDonald, J.S. (1999) The Global Land One-kilometer

586 Base Elevation (GLOBE) Digital Elevation Model, Version 1.0. . In, National
587 Oceanic and Atmospheric Administration, National Geophysical Data Center,
588 325 Broadway, Boulder, Colorado 80303, U.S.A. .
589 Hickman, C.P., Keen, S.L., Eisenhour, D.J., Ober, W.C., Larson, A., Ober, C.W. &
590 I'Anson, H. (2019) *Integrated Principles of Zoology*. McGraw-Hill Education.
591 Hortal, J., Triantis, K.A., Meiri, S., Thébault, E. & Sfenthourakis, S. (2009) Island
592 species richness increases with habitat diversity. *The American Naturalist*, **174**,
593 E205-E217.
594 IUCN (2022) The IUCN Red List of Threatened Species. Version 2022-2.
595 https://www.iucnredlist.org. In:
596 Jenkins, C.N., Pimm, S.L. & Joppa, L.N. (2013) Global patterns of terrestrial vertebrate
597 diversity and conservation. *Proceedings of the National Academy of Sciences*,
598 **110**, E2602-E2610.
599 Jetz, W., Sekercioglu, C.H. & Watson, J.E.M. (2008) Ecological correlates and
600 conservation implications of overestimating species geographic ranges.
601 *Conservation Biology*, **22**, 110-119.
602 Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers, A.O. (2012) The global
603 diversity of birds in space and time. *Nature*, **491**, 444-448.
604 Jung, M., Dahal, P.R., Butchart, S.H.M., Donald, P.F., De Lamo, X., Lesiv, M., . . .
605 Visconti, P. (2020) A global map of terrestrial habitat types. *Scientific Data*, **7**,
606 256.
607 Jung, M., Arnell, A., de Lamo, X., Garcia-Rangel, S., Lewis, M., Mark, J., . . . Visconti,

608 P. (2021) Areas of global importance for conserving terrestrial biodiversity,
609 carbon and water. *Nature Ecology & Evolution*, **5**, 1499-+.
610 Karr, J.R. (1971) Structure of avian communities in selected Panama and Illinois
611 habitats. *Ecological Monographs*, **41**, 207-233.
612 Keith, D.A., Ferrer-Paris, J.R., Nicholson, E., Bishop, M.J., Polidoro, B.A., Ramirez-
613 Llodra, E., . . . Kingsford, R.T. (2022) A function-based typology for Earth's
614 ecosystems. *Nature*, **610**, 513-+.
615 Lami, F., Bartomeus, I., Nardi, D., Beduschi, T., Boscutti, F., Pantini, P., . . . Marini, L.
616 (2021) Species–habitat networks elucidate landscape effects on habitat
617 specialisation of natural enemies and pollinators. *Ecology Letters*, **24**, 288-297.
618 Landi, P., Minoarivelo, H.O., Brännström, Å., Hui, C. & Dieckmann, U. (2018)
619 Complexity and stability of ecological networks: a review of the theory.
620 *Population Ecology*, **60**, 319-345.
621 Liu, H., Liu, Z., Zhang, M., Bascompte, J., He, F. & Chu, C. (2021) Geographic
622 variation in the robustness of pollination networks is mediated by modularity.
623 *Global Ecology and Biogeography*, **30**, 1447-1460.
624 MacArthur, R., Recher, H. & Cody, M. (1966) On the relation between habitat selection
625 and species diversity. *The American Naturalist*, **100**, 319-332.
626 Mariani, S., Ren, Z.-M., Bascompte, J. & Tessone, C.J. (2019) Nestedness in complex
627 networks: observation, emergence, and implications. *Physics Reports*, **813**, 1-
628 90.
629 McCain, C.M. (2009) Global analysis of bird elevational diversity. *Global Ecology and*

Biogeography, **18**, 346-360.

McGill, B.J. (2019) The what, how and why of doing macroecology. *Global Ecology and Biogeography*, **28**, 6-17.

Nardi, D., Lami, F., Pantini, P. & Marini, L. (2019) Using species-habitat networks to inform agricultural landscape management for spiders. *Biological Conservation*, **239**, 108275.

Newman, M.E. (2006) Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 8577-8582.

Nicodemus, K.K., Malley, J.D., Strobl, C. & Ziegler, A. (2010) The behaviour of random forest permutation-based variable importance measures under predictor correlation. *Bmc Bioinformatics*, **11**

Olesen, J.M., Bascompte, J., Dupont, Y.L. & Jordano, P. (2007) The modularity of pollination networks. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19891-19896.

Palmeirim, A.F., Emer, C., Benchimol, M., Storck-Tonon, D., Bueno, A.S. & Peres, C.A. (2022) Emergent properties of species-habitat networks in an insular forest landscape. *Science Advances*, **8**

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., . . . Sexton, J.O. (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *science*, **344**, 1246752.

Potts, S.G., Biesmeijer, J.C., Kremen, C., Neumann, P., Schweiger, O. & Kunin, W.E.

(2010) Global pollinator declines: trends, impacts and drivers. *Trends in ecology & evolution*, **25**, 345-353.

Powers, R.P. & Jetz, W. (2019) Global habitat loss and extinction risk of terrestrial vertebrates under future land-use-change scenarios. *Nature Climate Change*, **9**, 323-329.

Prasad, A.M., Iverson, L.R. & Liaw, A. (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, **9**, 181-199.

R-Core-Team (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raffaelli, D. (2004) How extinction patterns affect ecosystems. *Science*, **306**, 1141-1142.

Rahbek, C. (2005) The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology Letters*, **8**, 224-239.

Rivas-Salvador, J., Hořák, D. & Reif, J. (2019) Spatial patterns in habitat specialization of European bird communities. *Ecological Indicators*, **105**, 57-69.

Roll, U., Feldman, A., Novosolov, M., Allison, A., Bauer, A.M., Bernard, R., . . . Meiri, S. (2017) The global distribution of tetrapods reveals a need for targeted reptile conservation. *Nature Ecology & Evolution*, **1**, 1677-1682.

Rosenzweig, M.L., Turner, W.R., Cox, J.G. & Ricketts, T.H. (2003) Estimating diversity in unsampled habitats of a biogeographical province. *Conservation*

674 *Biology*, **17**, 864-874.
675 Sebastián-González, E., Dalsgaard, B., Sandel, B. & Guimarães Jr, P.R. (2015)
676 Macroecological trends in nestedness and modularity of seed-dispersal
677 networks: human impact matters. *Global Ecology and Biogeography*, **24**, 293-
678 303.
679 Seoane, J., Laiolo, P. & Obeso, J.R. (2017) Abundance leads to more species,
680 particularly in complex habitats: a test of the increased population size
681 hypotheses in bird communities. *Journal of Biogeography*, **44**, 556-566.
682 Thébault, E. (2013) Identifying compartments in presenceabsence matrices and
683 bipartite networks: insights into modularity measures. *Journal of Biogeography*,
684 **40**, 759-768.
685 Thébault, E. & Fontaine, C. (2010) Stability of ecological communities and the
686 architecture of mutualistic and trophic networks. *Science*, **329**, 853-856.
687 Tulloch, A.I.T., Barnes, M.D., Ringma, J., Fuller, R.A. & Watson, J.E.M. (2016)
688 Understanding the importance of small patches of habitat for conservation.
689 *Journal of Applied Ecology*, **53**, 418-429.
690 Wright, M.N. & Ziegler, A. (2017) ranger: A fast implementation of random forests for
691 high dimensional data in C++ and R. *Journal of Statistical Software*, **77**, 1 - 17.
692 Yodzis, P. (1980) The connectance of real ecosystems. *Nature*, **284**, 544-545.
693

694 **Data availability statement**

695 The data used to reproduce the results of the study are available in the database under

accession code in Figshare (https://figshare.com/s/4d7454f0ca87581a1f01). The distribution data of mammals and amphibians are available at https://www.iucnredlist.org/resources/spatial-data-download; The distribution data of reptiles are available at https://datadryad.org/stash/dataset/doi:10.5061/dryad.83s7k; The distribution data of birds are available at https://datazone.birdlife.org/species/requestdis; The data of species' suitable habitat types are available at https://www.iucnredlist.org; The map of terrestrial habitat types are available at https://zenodo.org/records/4058819;

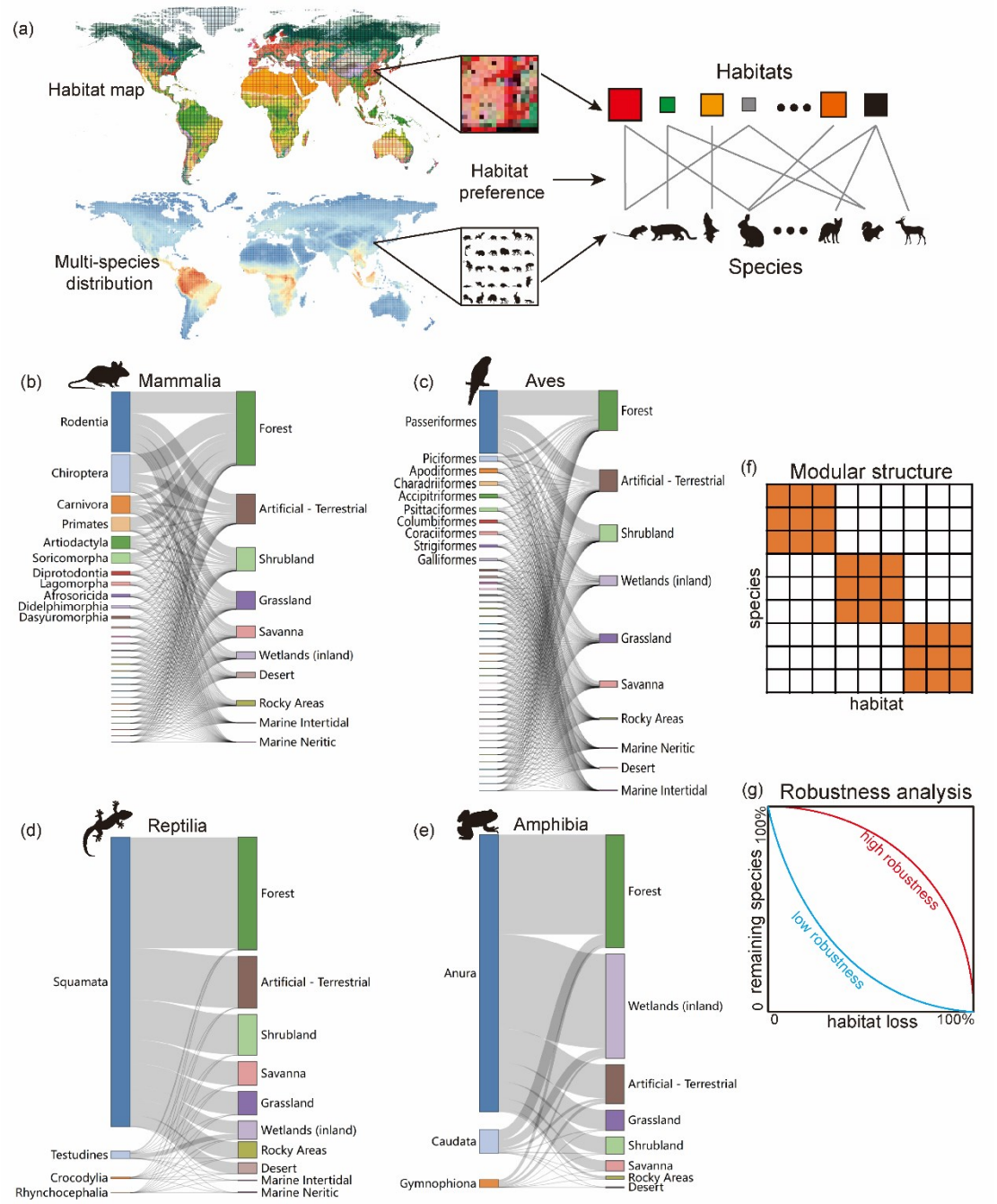
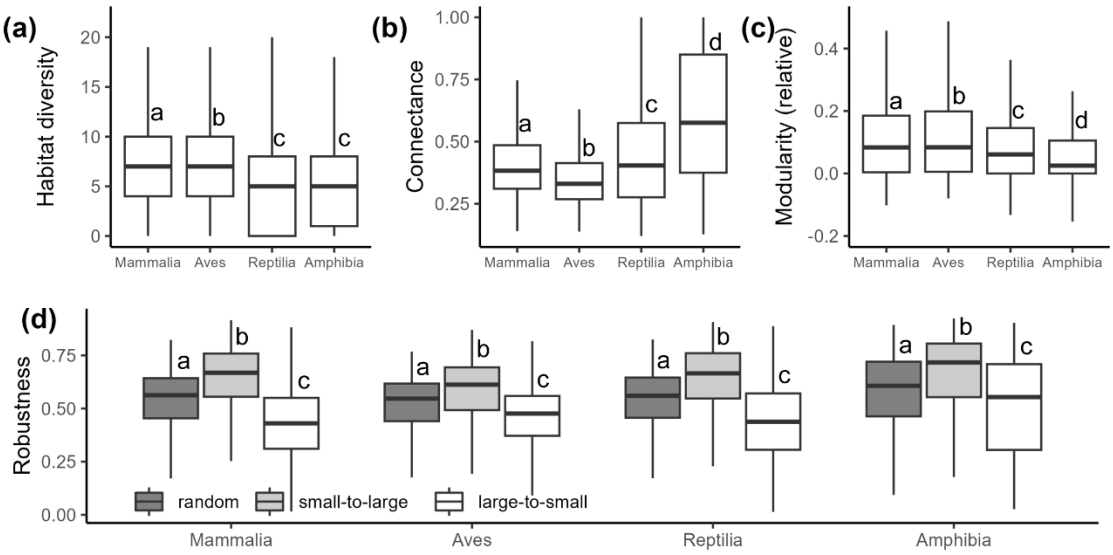


Fig. 1. Construction of species-habitat networks (a), global species-habitat meta-networks (b-e) and conceptual diagrams illustrating modularity (f) and robustness (g). The left-side nodes of each bipartite network represent the species (here summarized with orders), and the right-side nodes represent the habitats (level 1 habitat names shown).



712

713 Fig. 2. Comparison of structures(a-c) and robustness (d) of species-habitat networks

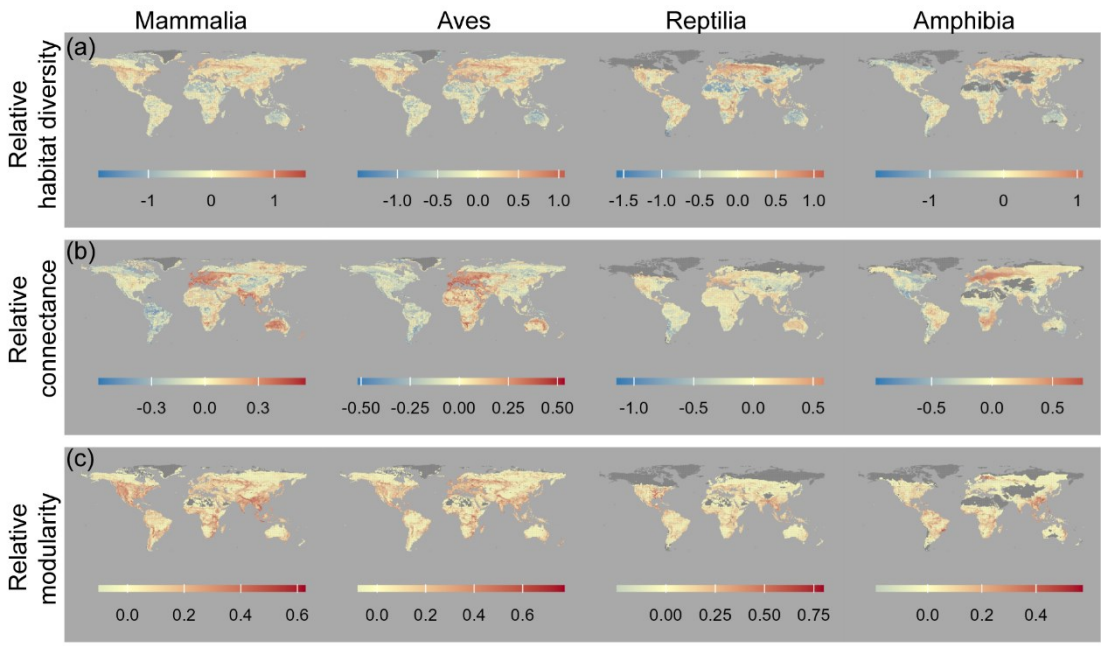
714 among four classes. The different low-case letters within plots indicate statistical

715 differences (P<0.05) among classes (a-c), and that indicate statistical differences

716 (P<0.05) among scenarios (d). Modularity is standardized by the “quasiswap” null

717 model.

718



719

720 Fig. 3. Spatial patterns of species-habitat network structures of four classes.

721 Modularity is standardized by the “quasiswap” null models. The dark grey regions

722 show empty values due to no species distribution or insufficient data for calculating

723 network metrics.

724

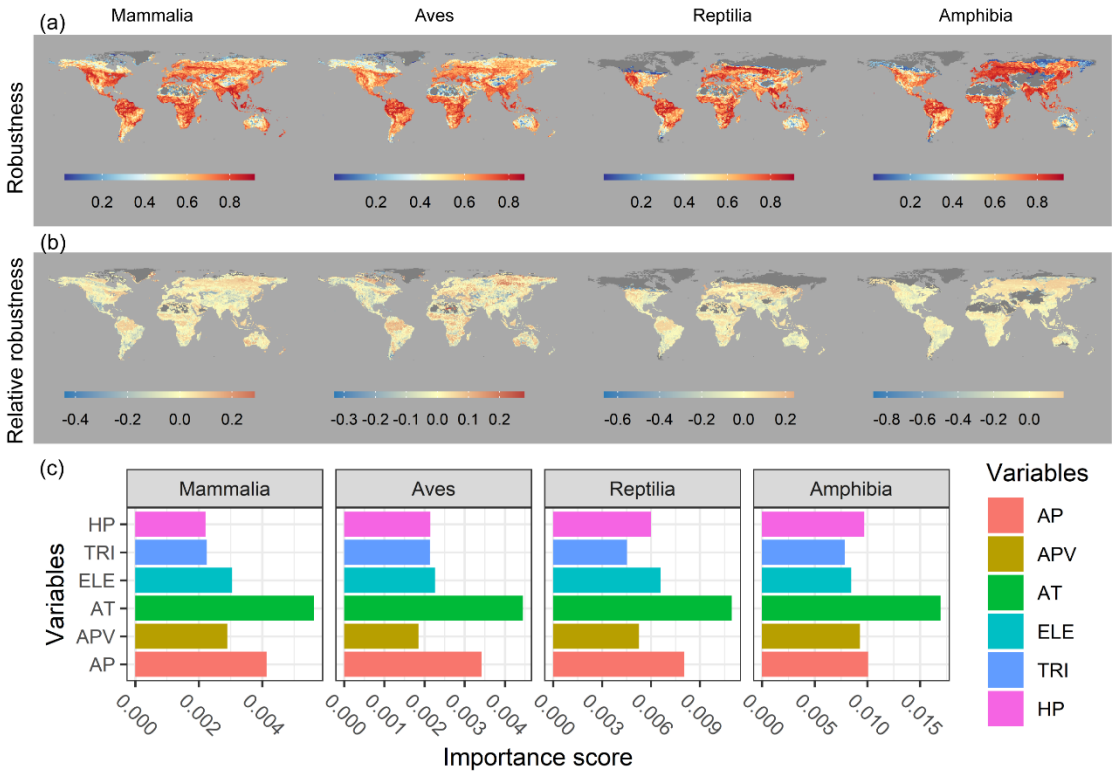
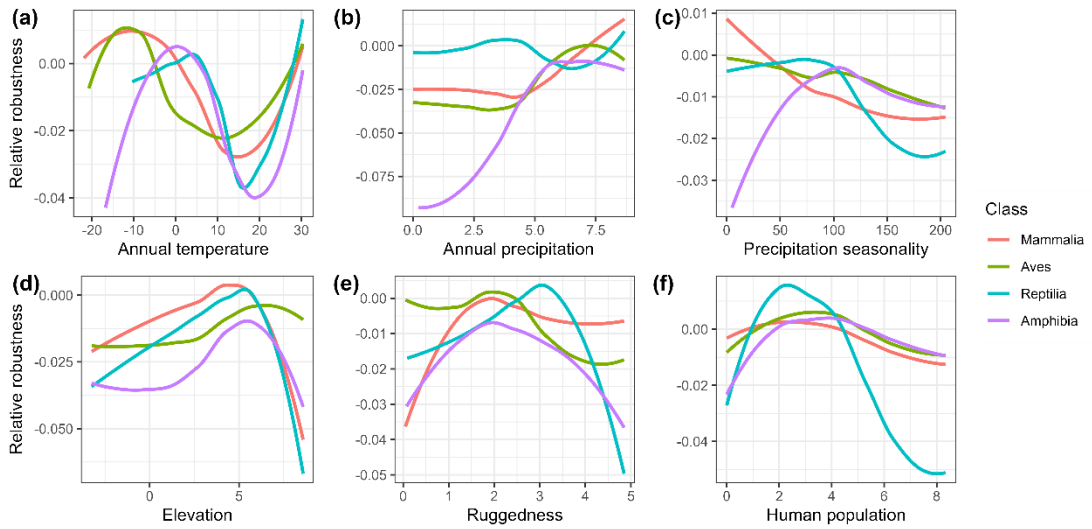


Fig. 4. Spatial patterns of network robustness (a) and relative robustness (b) from the small-to-large scenario and the importance of environmental variables (c) in affecting relative robustness from random forest models. AP, annual precipitation; APV, precipitation seasonality; AT, annual temperature; ELE, elevation; TRI, terrain ruggedness index; HP, human population density.

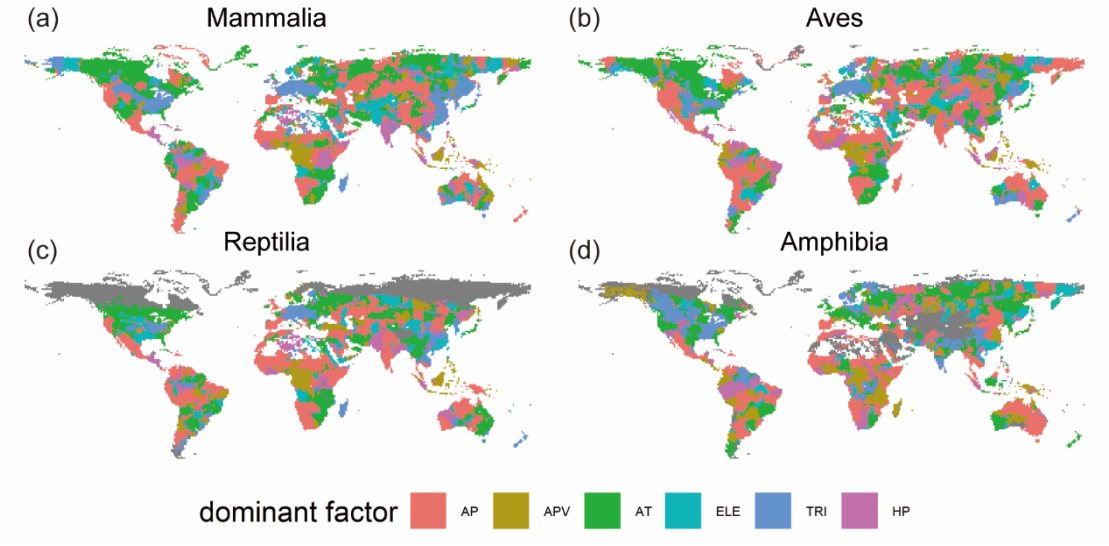


732

733 Fig. 5. The smoothed partial dependence plots showing the associations of

734 environmental variables with relative robustness from global random forest models.

735



736

737 Fig. 6. The spatial distributions of dominant factors for relative robustness (from the

738 small-to-large scenario) in four classes. AP, annual precipitation; APV, precipitation

739 seasonality; AT, annual temperature; ELE, elevation; TRI, terrain ruggedness index;

740 HP, human population density. The dark grey regions show empty values due to no

741 species distribution or insufficient data for calculating network metrics.

742