

ECOGRAPHY

Research

Biogeography of bird and mammal trophic structures

Manuel Mendoza and Miguel B. Araujo

EDITOR'S
CHOICE

M. Mendoza (<https://orcid.org/0000-0002-2143-8138>) ✉ (mmendoza@mmcn.csic.es) and M. B. Araujo (<https://orcid.org/0000-0002-5107-7265>) ✉ (maraujo@mmcn.csic.es), Depto de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. MBA also at: Rui Nabeiro Biodiversity Chair, MED Inst., Évora, Portugal.

Ecography

2022: e06289

doi: 10.1111/ecog.06289

Subject Editor:

Catherine H. Graham

Editor-in-Chief:

Jens-Christian Svenning

Accepted 16 February 2022



Does climate determine the trophic organization of communities around the world? A recent study showed that a limited number of community trophic structures emerge when co-occurrence of trophic guilds among large mammals is examined globally. We ask whether the pattern is general across all terrestrial mammals ($n = 5272$) and birds ($n = 9993$). We found that the six community-trophic structures previously identified with large mammals are largely maintained when all mammals and birds are examined, both together and separately, and that bioclimatic variables, including net primary productivity (NPP), are strongly related to variation in the geographical boundaries of community trophic structures. We argue that results are consistent with the view that trophic communities are self-organized structures optimizing energy flows, and that climate likely acts as the main control parameter by modulating the amount of solar energy available for conversion by plants and percolated through food webs across trophic communities. Gradual changes in climate parameters would thus be expected to trigger abrupt changes in energy flows resulting from phase transitions (tipping points) between different dynamical stable states. We expect future research to examine if our results are general across organisms, ecosystems, scales and methodologies, and whether inferences rooted in complex systems theory are supported. The emergence of general patterns in the functional properties of animal communities at broad scales supports the emergence of food-web biogeography as a sub-discipline of biogeography focused on the analysis of the geographical distributions of trophic relationships among organisms.

Keywords: biogeography, climate change, complex systems theory, food webs, energy optimization, net primary productivity, trophic structures

Introduction

Biogeography examines distributions of species and communities at broad geographical scales of extent. It is often assumed that the effects of biotic interactions on the distributions of species constitute a second-order effect (also known as the Eltonian noise; Soberón and Nakamura 2009), with the relationship between distributions and abiotic factors, such as climate, being more important (Whittaker et al. 2001, Pearson and



www.ecography.org

© 2022 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

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

Dawson 2003). These views have been challenged by authors examining the potential effects of biotic interactions on species range dynamics using theoretical (Mendoza et al. 2004, Gravel et al. 2011, Araújo and Rozenfeld 2013, Post 2013), empirical (Araújo and Luoto 2007, Baselga and Araújo 2010, Meier et al. 2010) and experimental settings (Davis et al. 1998, Callaway et al. 2002). However, the extent to which biotic interactions constrain large-scale biodiversity patterns is still a matter of ongoing debate: are broad scale geographical distributions of species largely driven by climate? Alternatively, are such distributions markedly constrained by biotic interactions?

Mendoza and Araújo (2019) contributed to this debate by asking how climate variables would determine the worldwide distribution of trophic structures, describing how species are distributed among feeding groups, and how these groups relate to one another. Using data on the geographical distributions and feeding preferences of large terrestrial mammals ($n=689$), they detected six trophic structures broadly matching the distribution of climate. They proposed that if the regularities detected with community trophic structure for large mammals were general across sets of organisms, a reasonable interpretation would be that climate does not just limit the distributions of species (Peterson et al. 2011), but also the types of interactions that are possible at any given location. Climate could thus be seen as acting as a boundary condition within which specific sets of trophic interactions co-exist regardless of the identity of the species available for colonization within the species pool (Cornell and Harrison 2014). Should the assumption be correct, a trophic community in a humid tropical region in Southeast Asia, for example, would be functionally equivalent to a trophic community in humid tropical areas of Africa, South America or northern Australia. To test the generality of these observations, we expand the analyses conducted by Mendoza and Araújo (2019) to all non-marine bird ($n=9993$) and terrestrial mammal ($n=5272$) species (total $n=15\,265$). Other vertebrate groups, such as amphibians and reptiles, were not included because there is insufficient available information on their feeding behaviour.

Background and predictions

The idea of communities being largely determined by climate and converging across areas with similar climates is inspired by the view of trophic interactions being the outcome of self-organized processes within communities (Solé and Bascompte 2012). Communities, like other complex dynamic systems, can be described by n state variables belonging to an n -dimensional space known as the state space. The structure of the communities can, in turn, be described as networks of energy flows whose nodes are groups of species exploiting a certain type of resources (i.e. trophic guilds) (Zhang 2009, Moore and De Ruiter 2012). Within a complex systems framework, the state variables describing the trophic community are the amount of energy processed by all individuals within each guild. Currently, we lack the tools to

conveniently measure energy flows across trophic community structures, especially for multiple species interactions across at large scales of extent and resolution. As an approximation to the problem one could ideally use abundance weighted by body mass as a measure the energy each guild processes. But since cell-based species abundance data across the world do not exist, we use the number of species within each trophic guild. The underlying assumption is that higher energy availability, which is related to greater amounts of resources, induce greater numbers of individuals within communities, consequently increasing species richness by allowing for greater numbers of species with viable populations. Empirical support for this assumption is mixed and still being discussed (Storch et al. 2018) but, given absence of global abundance data, the trophic structure of communities, defined by the number of species of each trophic guild, is used as a surrogate of the energy organization of that community (Mendoza and Araújo 2019).

In a dynamical system, all possible trajectories representing the motion of initial states under the system dynamics conform the phase portrait. An attractor is a region of the phase portrait towards which the dynamical system evolves for a set of initial states, which conforms its basin of attraction. System states that reach the vicinity of an attractor remain close to the attractor, even if they are slightly perturbed. Thus, although any configuration of the state variables describing the behaviour of the system could occur a priori, only some of them – the attractors – eventually emerge under the system dynamics. Trophic communities, as other dynamics systems, are expected to have stable configurations described as attractors of the energy transfer dynamics (Scheffer and Carpenter 2003, Mendoza et al. 2004, Moore and De Ruiter 2012). That is, trophic communities should not be formed by any combination of species or guilds, just but by those contributing to reaching a stable organization of energy flows.

The phase portrait topology, i.e. the number, location and shape of attractors and their basin of attraction, is determined by a set of p parameters known as control parameters, which define a p -dimensional space known as the parameter space. While the control parameters (e.g. climate variables) determine the emergent dynamics of the system (e.g. energetic organization of the food webs), thus affecting the individual behaviour of the state variables (e.g. energy processed by the trophic guilds), they are not affected by such organization. This does not imply that biota does not affect climate. It does but, as an approximation, a state variable can be treated as a control parameter when the time scale at which it acts is much larger (e.g. evolutionary changes in ecological surveys), or also when its influence on the dynamics of the system is much larger than the one it has on it (e.g. mean temperature or precipitation on the energy organization of communities).

In terrestrial ecosystems it has been known for over 200 hundred years that the availability of energy and water throughout the year largely determines the amount and type of vegetation that can grow (Humboldt 1807, Holdridge 1947) and the production rate of different plant resources for animals (Hawkins et al. 2003). Thus, it is reasonable to expect

that climate variables functioning as key control parameters for vegetation would also act as control parameters of trophic communities (Currie et al. 2004) even if lagged responses to historical climates changes might occasionally be expected, chiefly among dispersal-limited species (Araújo et al. 2008) or clades limited by niche conservatism (Romdal et al. 2013, Bennet et al. 2021). In such cases, individual species–climate relationships can be stronger with historical climate than with contemporary climate even if non-equilibrium causing species to be absent from otherwise suitable climates (Araújo and Pearson 2005, García-Valdés et al. 2013) might be compensated locally by functional equivalence among species (Loreau 2004, Hubell 2005).

Small changes in control parameters of complex systems can result in small variations in the location and shape of attractors and their basin of attraction. In some cases, a small change in some parameters can trigger drastic changes in the phase portrait topology. These drastic changes, known as phase transitions or tipping points, can lead to alterations in the emerging dynamics and organization of the system. It can be the case, for example, of the transitions between vegetation dominated by different functional types such as biomes (De Angelis 2021). As a consequence of phase transitions, qualitative changes in the trophic structure of terrestrial communities have also been detected in model systems (Mendoza et al. 2004, Mendoza 2008, Boettiger and Batt 2020).

Demonstrating that gradual changes in the control parameters governing community dynamics can lead to emergence of well-defined groups separated from each other by abrupt transitions is far from trivial, because observations of these phenomena are difficult under fully controlled conditions. Previous studies have used indicators at local to regional geographical scales using field data (Scheffer and Carpenter 2003). They included examinations of: 1) jumps in time series, as between eutrophic and dystrophic states of lake ecosystems associated with distinct differences in phosphorus and organic carbon input (Carpenter and Pace 2013) or the collapse of Saharan vegetation triggered by the gradual orbital increases in summer season insolation (Demenocal et al. 2000); 2) spatial patterns, such as patchwork mosaics of barren areas and kelp beds that characterize many kelp forest ecosystems (Konar and Estes 2003), striking geometric patterns in the vegetation of extensive areas of arid or semi-arid Africa (Couteron and Lejeune 2001) or isolated vegetation patches observed in nutrient-poor territories of South America and West Africa (Lejeune et al. 2002); and 3) the emergence of discrete units as lakes with scarce or abundant submerged macrophytes (Van Geest et al. 2003).

If gradual changes in climate parameters trigger abrupt changes in energy flows resulting from phase transitions between different dynamical stable states (Beisner et al. 2003, Scheffer and Carpenter 2003, Mendoza et al. 2004, Moore and De Ruiter 2012), distinguishable trophic structures should emerge at global scale (Mendoza and Araújo 2019). Our first prediction is that communities continuously distributed in a geographical space should not be distributed continuously in the trophic space, defined by the number of

species of each trophic guild – a surrogate of their energy organization – rather forming discrete groups. Assuming that climate variables are the main control parameters of trophic terrestrial communities at broad scales, via their direct effects on net primary productivity (NPP) and broad vegetation types (von Humboldt and Bonpland 1807, Holdridge 1947, Woodward 1987), our second prediction is that regions with similar climatic conditions should feature similar trophic structures regardless of where they occur in the planet.

Methods

Data

Three sources of geographical data were extracted and plotted in a world terrestrial $1 \times 1^\circ$ grid system: 1) global distributional ranges of non-marine mammal and bird species; 2) bioclimatic variables; and 3) net primary productivity. The global species distributions were derived from IUCN Global Assessment distributional data for native ranges (IUCN 2014). Specific occurrences in grid cells were used to produce a presence/absence matrix with names of 9993 non-marine birds and 5272 terrestrial mammals (15 265 species) as columns and 18 418 $1 \times 1^\circ$ grid cells as rows.

The feeding habits of the bird and mammal species were obtained from the global species-level compilation of key attributes published by Wilman et al. (2014). No records for species occurring in Antarctica were available.

Bioclimatic data for the terrestrial surface of the Earth were obtained from WorldClim – Global Climate Data (Hijmans et al. 2005). We used the full set of 19 bioclimatic variables available in WorldClim for exploration of bioclimatic correlates of biological patterns inferred. The variables represent annual trends (e.g. mean annual temperature, annual precipitation), seasonality (e.g. annual range in temperature and precipitation) and extreme or limiting climatic factors (e.g. temperature of the coldest and warmest month, and precipitation of the wet and dry quarters).

Net primary productivity (NPP) data at 0.25×0.25 decimal degree were obtained from Imhoff et al. (2004), and then resampled to the $1 \times 1^\circ$ degree grid system used. It represents the total amount of carbon absorbed by land plants every year that is fixed in plant structures. NPP is the primary trophic source of energy for consumers and is measured in grams of carbon per grid cell.

In order to investigate whether community trophic structures are not just affected by cell-based climate conditions, but also by broader regional patterns, we also investigated the relationship between trophic structure membership and averaged climate values across neighbour cells within a radius of $1\text{--}8^\circ$ (averaged values). Thus, from each variable we derived 9 metrics (Garcia et al. 2014).

Identification of trophic guilds

In the global species-level compilation published by Wilman et al. (2014), trophic resources are classified into

10 categories shared by terrestrial mammals and non-marine birds (Fig. 1). The trophic profile of each species is obtained from the estimated percentage of each type of resource in their diet (similar procedure was undertaken by Mendoza and Araújo 2019). The result is a matrix with the 10 trophic resources categories as columns, 15 265 species of birds and mammals as rows, and values representing the estimated percentage of each type of resource. The trophic profile of each species is thus a point in a 10-dimensional 'trophic space' defined by the estimated percentage of each type of resource in their diet (a vector of dimension 10). This 10-dimensional space, which includes all mammal and bird species together, will be referred to here as 'species-level trophic space', as opposed to the 'community-level trophic space' described below.

Trophic guilds are commonly defined based on arbitrary thresholds regarding the percentage of each type of resource in the diet of the species (Janis 1990, Wilman et al. 2014). Consequently, species located next to either side of the threshold are classified into different groups, despite consuming very similar types of resources. To avoid such arbitrariness and sensitivity to threshold choice, trophic guilds were obtained with c-means clustering (Dunn 1973, Bezdek 1981), on the basis of the Euclidean distance between the 15 265 species in the 10-dimensional 'species-level trophic space' described above. Clustering analysis takes advantage of the less densely populated regions of this 10-dimensional trophic space to make the divisions, thus minimizing the number of species having very similar types of food and being classified into

different trophic groups. The number of clusters was selected based on the AMD_i curve (Supporting information).

Identification of community trophic structures

A $1 \times 1^\circ$ grid cell surface is about 12 300 km² at the equator. Such cells will often include several types of communities. However, for convenience, and assuming that differences within cells will be smaller than differences between cells located in different biogeographic regions, we considered species found in individual cells as belonging to the trophic community of that cell (and referred to henceforth as 'community'). Of course, this is a simplification because species within a cell might pertain to different trophic structures (e.g. a Savannah crossed by river surrounded by lush tropical vegetation will have distinct trophic structures in each of the two habitats). Such assumptions of homogeneity within sampling units are common and inevitable in regionalization exercises at both global (Holt et al. 2013) and regional levels (Rodrigues et al. 2015).

As a starting point, we assigned the 15 265 terrestrial mammal and non-marine bird species to their correspondent guild and then counted the number of species of each guild within each cell. The result is a matrix with the 9 trophic guilds as columns, 18 418 communities as rows, and values representing numbers of species. The trophic profile of every community is thus a point in a 9-dimensional 'trophic space' defined by the number of species from each trophic guild (a vector of dimension 9). This 9-dimensional space

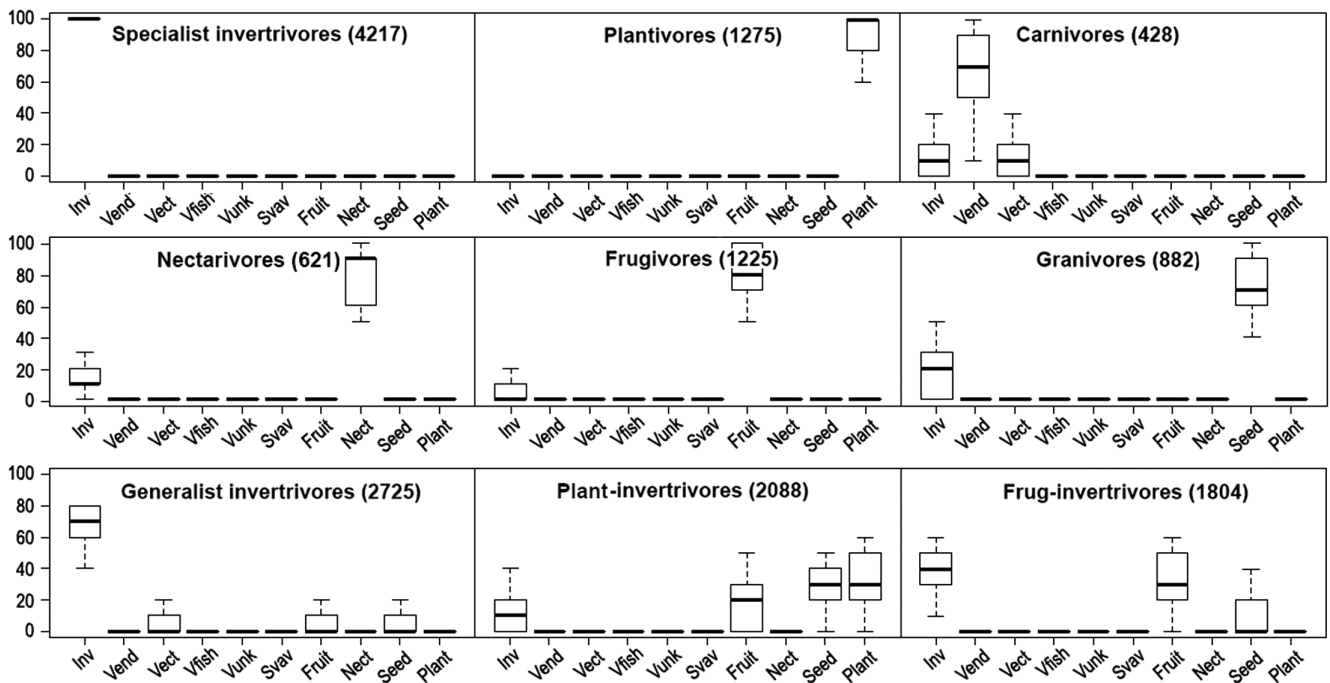


Figure 1. Distribution of the percentages (y-axis) of the ten types of trophic resources (x-axis) consumed by species across each one of the nine trophic guilds identified with c-means clustering. Trophic resources: invertebrates, *Inv*; vertebrate endotherms, *Vend*; vertebrate ectotherms, *Vect*; fish, *Vfish*; vertebrates in general, *Vunk*; carrion, *Svav*; fruit and drupes, *Fruit*; nectar and pollen, *Nect*; seed and grains, *Seed*; other plant material and fungi, *Plant* (see for detailed description Wilman et al. 2014).

will be referred to here as ‘community-level trophic space’, as opposed to the 10-dimensional ‘species-level trophic space’ described above.

Our first prediction is that communities, even if continuously distributed in a geographical space, will not be continuously distributed in the ‘community-level trophic space’, rather forming well-defined groups separated by empty or sparsely populated regions. A multidimensional space cannot be represented on a two-dimensional surface so, in order to test this prediction, we developed the average mean decrease analysis (‘AMD analysis’, Supporting information). The trophic structures identified were represented with boxplots showing the distribution of the numbers of species within nine trophic guilds.

The same analyses were also performed for mammals ($n=5272$) and birds ($n=9993$), independently, but not to test the first prediction with them, but to test the generality of the community trophic structures found, comparing them with each other, as well as with the structures identified combining both groups ($n=15\,265$) and those previously found for large terrestrial mammals ($n=689$, Mendoza and Araújo 2019).

Climatic modelling of community trophic structures

The second prediction proposes that regions with similar climatic conditions feature the same trophic structure regardless of their location in the planet. The relationship between trophic structure and both climate and NPP was modelled using two different machine learning techniques: Random Forests (Breiman 2001) with the package *RandomForest* (Liaw and Wiener 2002), and evolutionary learning of globally optimal classification trees (Krętownski and Grześ 2005) with the package *evtree* (Grubinger et al. 2017). Random forests were selected for their extraordinary capacity to deal with complex relationships, decorrelating highly correlated variables and taking advantage of much of the information contained in them (Boulesteix et al. 2012). It is based on hundreds of trees, which can hardly be analyzed to see how the predictor variables interact with each other to determine the behavior of the target variable (the type of trophic structure in this case). A simple tree, however, especially if optimized with evolutionary algorithms (such as the one used by the *evtree* package), can reveal some of the most important interactions, albeit in a simplified form.

Like other predictive models, the *RandomForest* package can provide predictions for out-of-the-bag (OOB) samples not used for training the algorithms (often referred to as ‘projection’ when the prediction domain is spatially uncorrelated with the model training domain, Araújo et al. 2019). When OOB samples are selected at random, they are inevitably similar to those used for training the model. This can cause a problem for projections since the similarity between training data and test data violates the statistical independence between the two sets owing to spatial autocorrelation. The result typically implies inflated assessments of models predictive performance (Araújo et al. 2005). To

limit the effects of spatial autocorrelation on the predictions, the algorithms were tested on previously excluded regions (Dormann et al. 2007). Firstly, we divided the globe into 36 bins of 10° longitude (from 180° West longitude to 180° East longitude). Secondly, we fitted the models with cells from all bins except one, which is used for testing. This procedure of 36-fold cross-validation is repeated so that all bins are used for testing once. Agreement between predicted and observed trophic structures on testing bins is measured with the kappa coefficient, which quantifies the overlap expected by chance (Carletta 1996), and a kappa value is thus obtained for all cross-validated model runs (see for similar approach Madon et al. 2013).

The importance of the predictor variables for the development of random forests was obtained from the average decrease of the Gini coefficient (MDG), which is based on the contribution that each of them makes to increase the homogeneity of the nodes when incorporated to the trees from which the random forests are constructed (Strobl et al. 2008).

Coastal communities, identified by zero distance to the sea, were excluded because energy input from the sea is expected to affect the energy organization of coastal communities independently of the climate. This left 15 370 out of 18 418 communities. NPP values were not available for some high-latitude regions, e.g. most of Greenland, so these regions were also not used in our study. The final number of cells used for both climate and NPP modeling was 14 563.

Results

Trophic guilds

We selected nine major trophic guilds (Supporting information, Fig. 1). Specialist invertivores and plantivores are considered specialist groups as they consume, on average, 100% of either invertebrates or plants. Carnivores, nectarivores, frugivores and granivores are well-defined trophic guilds, although they also consume to a small extent, invertebrates. The remaining 3 groups are generalist invertivores (as opposed to the specialist invertivores), plant-invertivores and frug-invertivores. Invertebrates are a significant component of the diet in these three groups. Some trophic resources, such as fish (Vfish), vertebrates in general (Vunk) or carrion (Scav) represent such low percentage of the total consumption that they played a small role in the formation of the trophic guilds.

Community trophic structures

The curve resulting from the application of the AMD_i analysis to the 18 418 world terrestrial grid cells describing the 9-dimensional community-level trophic space obtained with mammals and birds together showed a peak for 6 user-defined clusters (Fig. 2e), indicating that there are six well-defined groups (Supporting information). These groups are

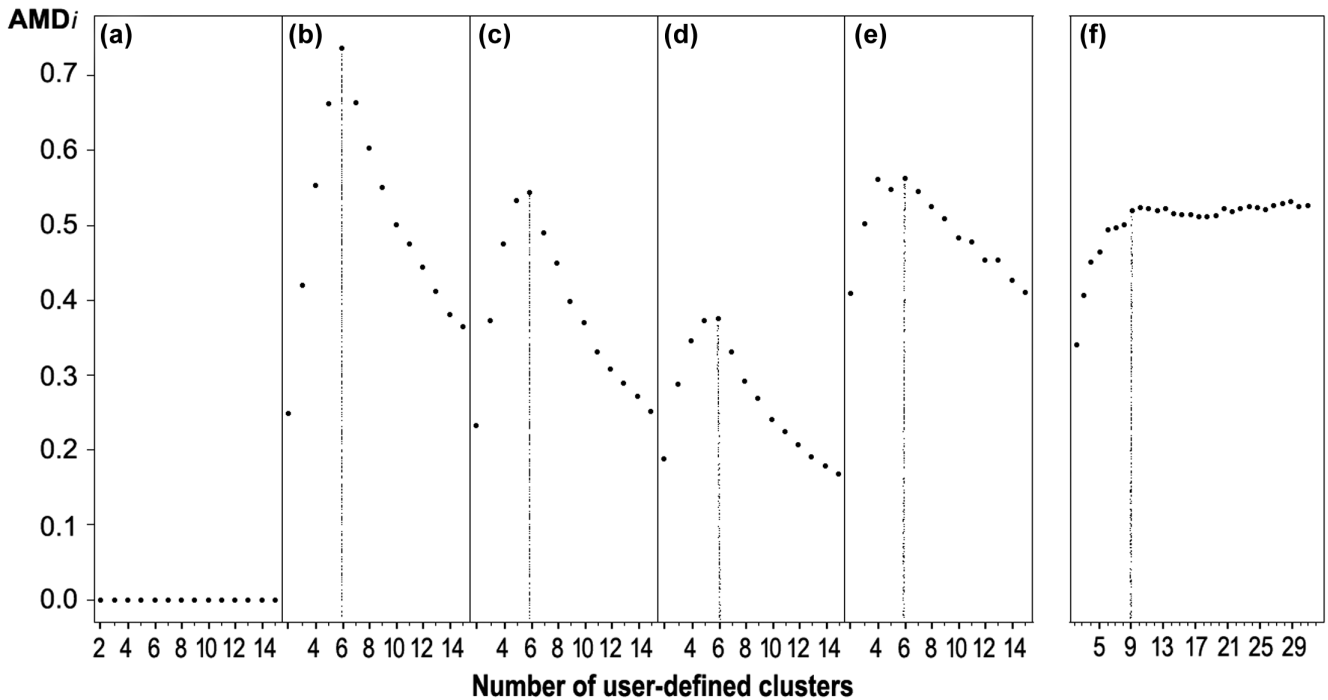


Figure 2. AMD_i curves: (a) obtained with 18 418 artificial samples randomly distributed in a 9-dimensional space of side 100; (b) 18 418 artificial samples in a 9-dimensional space of side 100, grouped in 6 clusters with standard deviation 5 (highest definition); (c) with standard deviation 10 (intermediate definition); (d) with standard deviation 15 (lowest definition); (e) 18 418 communities (cells) in the 9-dimensional community-level trophic space; (f) 15 265 species in the 10-dimensional species-level trophic space; y-axis: AMD_i values (Supporting information).

the six trophic structures, whose main characteristics are represented in Fig. 3. AMD_i reaches a value (0.56) comparable to that of Fig. 2c, which was obtained with the same number of samples forming six groups whose standard deviation was 10 (in a space of side 100). As expected, the communities from each group share a common pattern in the number and the relative proportion of coexisting guilds (Fig. 3).

The projection of the six community trophic structures on any of the 9-dimensional community-level trophic space (the number of species from each trophic guild) shows a significant overlap between them (Fig. 3). However, in the 9-dimensional space it is expected that the six trophic groups show an overlap equivalent to six clusters of samples with a standard deviation of only one tenth of the hypercube edge in which they are found.

The AMD_i curves obtained with mammals and birds independently also showed a peak for 6 user-defined groups. The overlap between the geographical distributions of the six trophic structures identified by both curves is high ($\kappa=0.73$, Fig. 4a and b respectively). The main differences between trophic structures identified for mammals and birds are on islands, which are differentially colonized by flying and non-flying animals, and also on the Indian Peninsula. The geographical limits of trophic structures 2 and 3 (dark blue and light blue respectively, Fig. 4a and b) are also mismatching in Asia. Finally, small areas in Europe with trophic

structure 4, present with mammals, are absent with birds (yellow, Fig. 4a and b).

Climatic modelling of community trophic structures

The geographical distribution of the six community trophic structures predicted by Random Forests using climatic predictors (Fig. 4d and e), closely matches the observed distribution (Fig. 4c). Using climatic variables as predictors at cell-level (Fig. 4d), the agreement between predicted and observed is substantial ($\kappa=0.73$, Table 1). When NPP is additionally included, the prediction capacity of the resulting model practically does not change ($\kappa=0.74$, Table 1), but NPP becomes the single most important variable (MDG, Table 1). Using smoothed climatic variables by averaging values across neighbour cells, the match between predicted (Fig. 4e) and observed structures (Fig. 4c) increases ($\kappa=0.81$, Table 1). When NPP is also included, κ increases only slightly ($\kappa=0.82$, Table 1), while still remaining the most important variable (MDG, Table 1). Using only local and averaged values of NPP, the agreement between predicted and observed decreases ($\kappa=0.54$, Table 1). A tree model including only averaged values of NPP (not shown here) explains the general geographical pattern of the global distribution of the six trophic structures around the world ($\kappa=0.59$, Fig. 4f). A tree model including cell-level

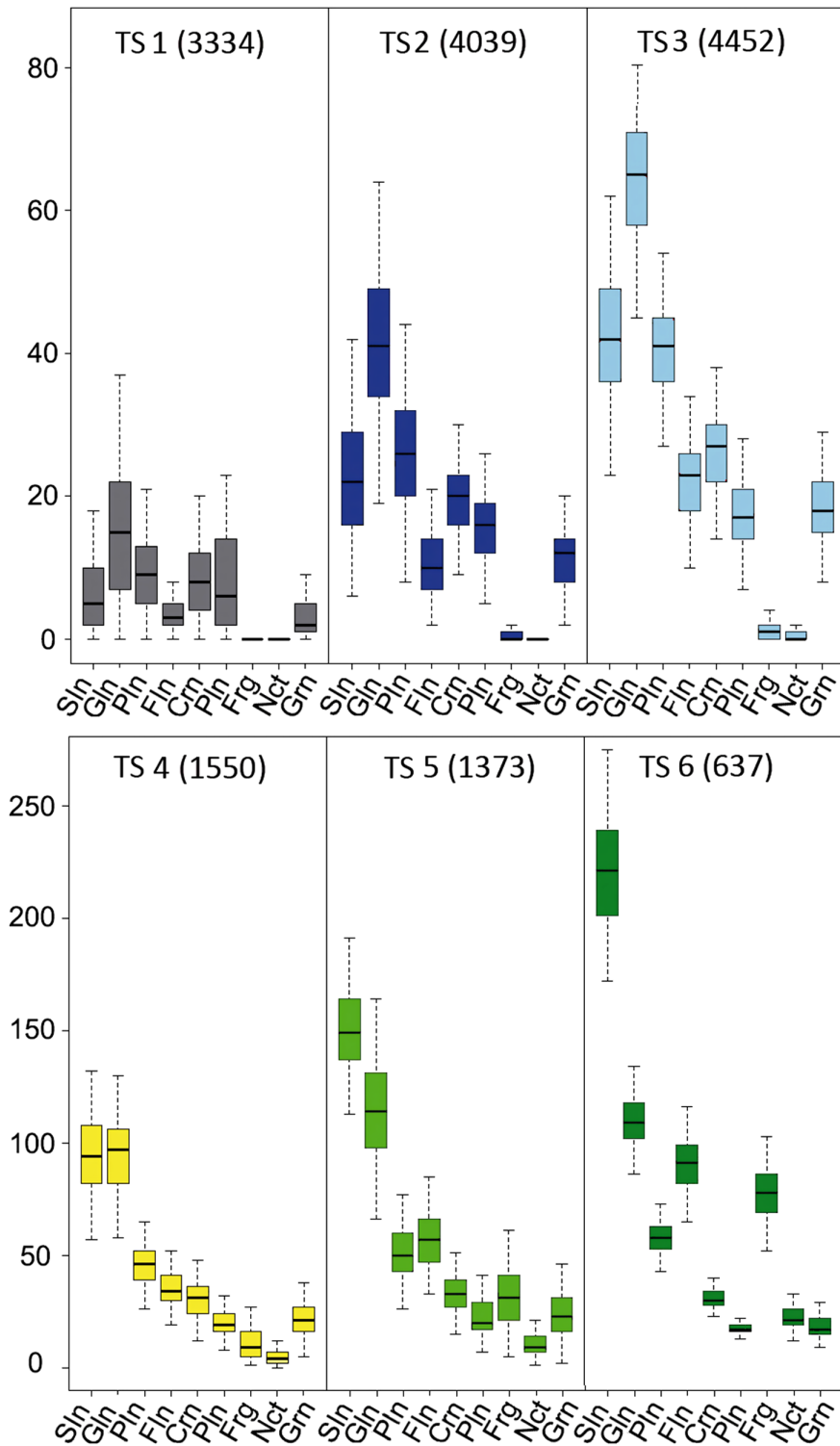


Figure 3. Representation of the six trophic structures identified with the AMD_i analyses. Frequency distribution of the numbers of species within nine trophic guilds across six trophic structures. Values within parentheses are number of cells. Trophic guilds are: specialist invertivores, *Slm*; generalist invertivores, *Gln*; plant-invertivores, *Pln*; frug-invertivores, *Fln*; carnivores, *Crm*; plantivores, *Plm*; frugivores, *Frg*; nectivores, *Nct*; granivores, *Gm*; note that the range of species numbers in the first three trophic structures (0–80) is lower than that of the last three (0–250).

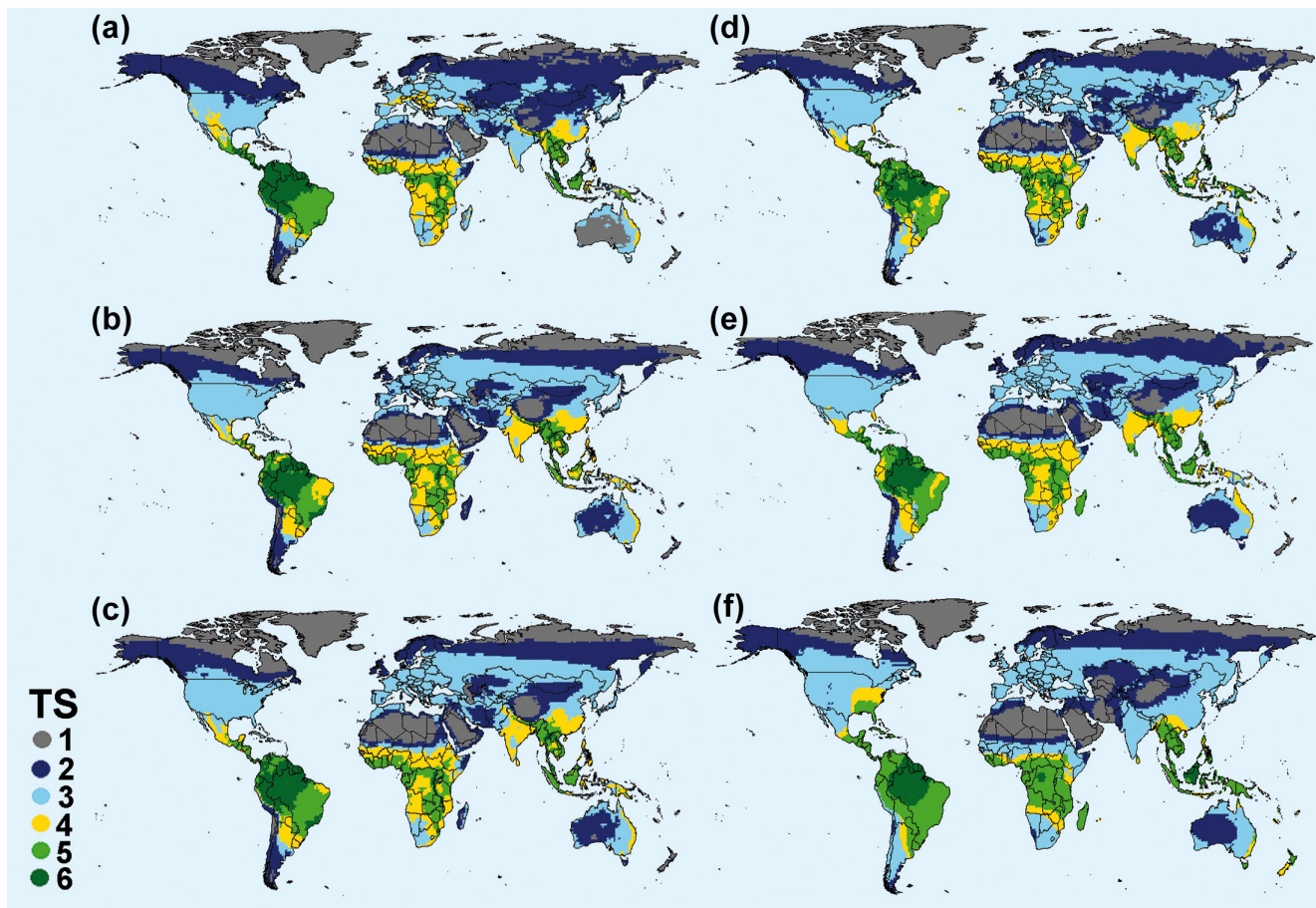


Figure 4. Geographical distribution of the six trophic structures with: (a) all mammal species (5272 species); (b) all bird species (9993 species); (c) both mammal and bird species together (15 265 species). Geographical distribution of the six trophic structures predicted (d) by random forests using climatic predictors, values at cell level ($\kappa=0.73$); (e) climatic predictors, values averaging across neighbour cells ($\kappa=0.81$); (f) by classification trees using only NPP values averaging across neighbour cells (tree not shown here, $\kappa=0.59$).

NPP and climatic variables (Fig. 5, $\kappa=0.64$) shows how climate modulates NPP to determine the trophic structure.

Discussion

Consistent with the two predictions made, we found that 1) well-defined community trophic structures emerge when examining trophic preferences of land mammals and birds around the world; and 2) the distribution of trophic

structures matches climatic gradients, areas sharing similar environmental characteristics also sharing similar community trophic structures regardless of the biogeographic region they occur. It is thus reasonable to interpret climate, namely the energy available for species, as a boundary condition within which trophic communities self-organize. In other words, climate should impose the limits on the types and quantities of resources available for first-order consumers, which in turn limit the types and quantities of resources available for second-order consumers, and so on.

Table 1. Predictor variable importance of community trophic structures using Random Forest models.

Predictors	Fig.	Kappa	MDG
Climate (cell level)	4d	0.73	BIO10, BIO13, BIO16
Climate and NPP (cell level)		0.74	NPP
Climate (cell level and averaged)	4e	0.81	BIO10 (averaged values)
Climate and NPP (cell level and averaged)		0.82	NPP (averaged values)
NPP (cell level and averaged)		0.54	NPP (averaged values)

The Kappa coefficient measures the agreement between observed and predicted trophic structures on test data (previously excluded longitudinal bins 10° width). The most important predictors are mean temperature of warmest quarter (BIO10), precipitation of wettest month (BIO13), precipitation of wettest quarter (BIO16) and net primary production (NPP).

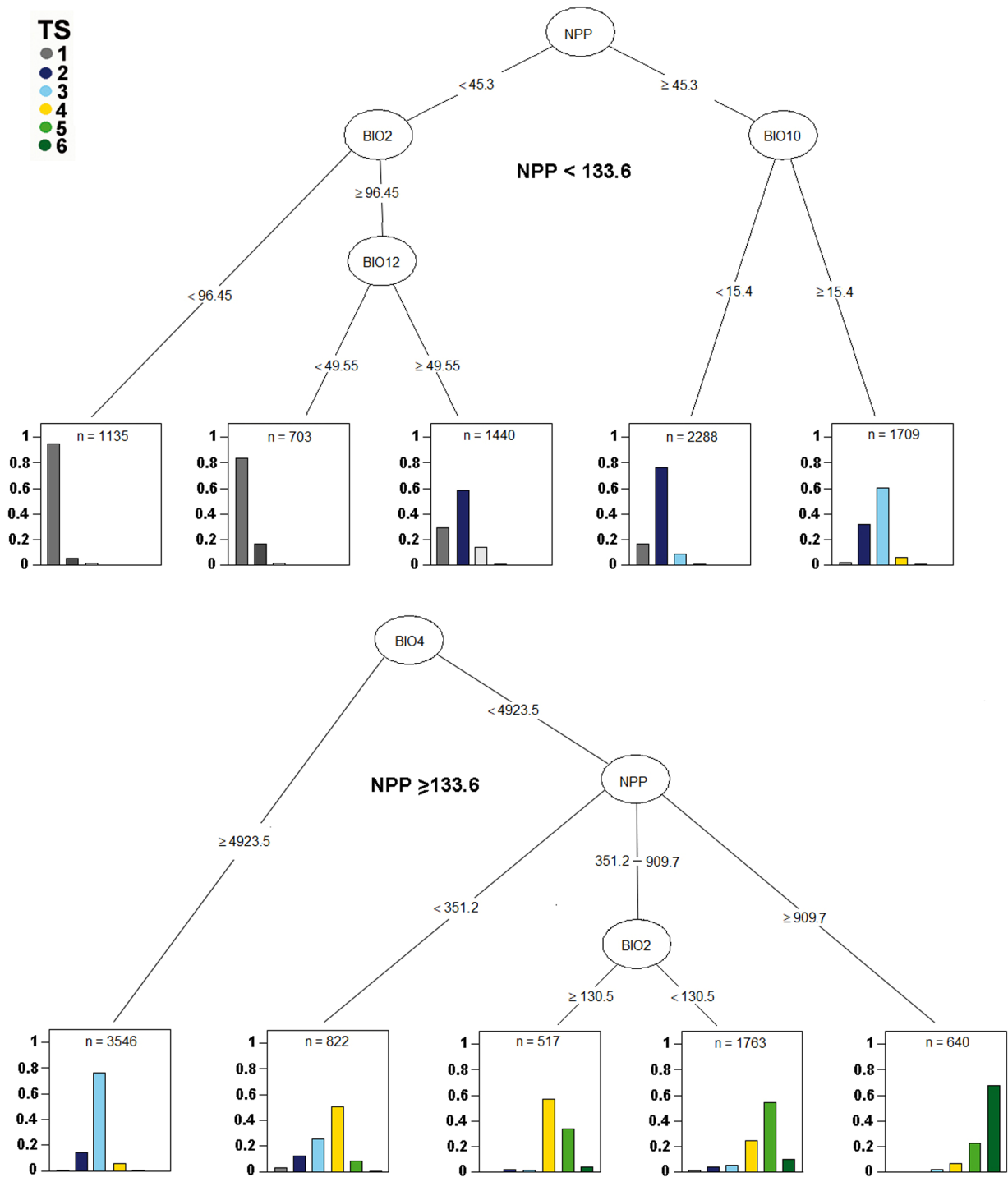


Figure 5. Classification tree model relating net primary production and climate with the six trophic structures ($\kappa = 0.64$). Its two main branches are shown separately for easy viewing. n indicates the total number of communities contained by that terminal node and y -axis the proportion of these communities corresponding to each type of trophic structure (TS). All the communities contained in each terminal node (lower) fulfil the conditions leading to it from the root node, which divides those communities whose NPP value < 133.6 from those in which the NPP value ≥ 133.6 . Variables: net primary productivity ($\times 10^9$), NPP; annual mean temperature, BIO1; mean diurnal range (mean of monthly (max temp – min temp)), BIO2; isothermality (BIO2/BIO7) ($\times 100$), BIO3; temperature seasonality (standard deviation $\times 100$), BIO4; max temperature of warmest month, BIO5; min temperature of coldest month, BIO6; temperature annual range (BIO5–BIO6), BIO7; mean temperature of wettest quarter, BIO8; mean temperature of driest quarter, BIO9; mean temperature of warmest quarter, BIO10; mean temperature of coldest quarter, BIO11; annual precipitation, BIO12; precipitation of wettest month, BIO13; precipitation of driest month, BIO14; precipitation seasonality (coefficient of variation), BIO15; precipitation of wettest quarter, BIO16; precipitation of driest quarter, BIO17; precipitation of warmest quarter, BIO18; precipitation of coldest quarter, BIO19.

Rather than simply conforming ‘point estimates of overlapping regional species distributions’ (Ricklefs 2008), animal communities at broad biogeographical scales can thus be seen as constrained by the sets of biotic interactions that are possible given the environmental characteristics of the area in which they occur (Mendoza and Araújo 2019). Species might be still be said to be doing ‘their own thing’ (acting independently from each other in order to maximize their fitness), as expected under the individualist viewpoint of community organization (Gleason 1926). However, their ‘degrees of freedom’ should be constrained by the limiting effects of climate on the critical physiological limits of consumers and the plant production rates affecting the quantity and quality of the biomass available for consumption. Within these limits, optimal foraging strategies (Pyke 1984, Parker and Smith 1990) might be invoked as one of the mechanisms driving the community self-organization leading to convergence in trophic structures across biogeographical regions with similar climates.

The proposed viewpoint offers an intermediate perspective between the two historically opposed perspectives on community organization (e.g. see for recent discussions Baselga and Araújo 2009, Loreau 2020) – the superorganism perspective, whereby strong assembly rules would determine the collective response of individuals and species within communities, and the individualistic perspective, whereby individual species would act independently of each other, and often idiosyncratically, to maximize the fitness of its individuals. Top down (and individualistic) perspectives propose that biotic interactions, namely predation, drive community dynamics and the ensuing patterns. There are local empirical examples supporting this view (Post et al. 1999) but at broad biogeographic levels the view is put to question by the overwhelming evidence of bottom–up abiotic environmental constraints controlling the distribution of major vegetation (community) types (Holdridge 1947). What our analysis reveals is that climate constraints do not simply affect individual species dynamics, as inferred from individualistically-inspired ecological niche theory underlying, for example, species distributions models (Peterson et al. 2011), but also constrains the types of trophic communities that can be expected in any particular location.

The existence of clearly defined animal communities at global level whose distributions match climatic gradients lends credence to the view that bottom–up processes control community dynamics. While ecologically-driven regularities in animal communities have not been detected in classic eco-regionalization exercises (Holt et al. 2013, Morrone 2015, Rodrigues et al. 2015), we believe the reason for this is that species co-occurrence patterns rather than their functional attributes were the focus. In such studies, faunal differentiation among regions exposed to similar climates emerges as a consequence of different evolutionary pathways caused by limited dispersal (Holt et al. 2013, Ficetola et al. 2017). For example, in a classic eco-regionalization, the faunas of Australasia are grouped differently from the faunas of the Nearctic even when they are exposed to similar climates

(Holt et al. 2013). However, when examining the composition of functional attributes within animal communities, instead of simply the species composition, regions with different evolutionary pathways are grouped together owing to their similar trophic structures. In other words, if communities are analysed not just by the composition of their species, which are affected by historical patterns of speciation, extinction and colonization, but also by their numbers and functional (trophic) properties, regularities emerge.

Although regions with the same climatic conditions tend to feature the same trophic structures, we also found that communities exposed to different climatic conditions could feature the same trophic structure. This is because net primary production (NPP), strongly covaries with trophic community structures. NPP responds non-linearly to energy and water and determines the amount of chemical energy processed by plants through photosynthesis (Cramer et al. 1999, Sitch et al. 2003). In our analysis, NPP is the variable best explaining geographic transitions among animal trophic structures, and areas with different climates can have equivalent trophic structures because NPP is similar. It is the case, for example, of both hot and cold deserts, or the Arctic regions – all sharing similar trophic structures (TS 1) and low inputs of energy. In deserts, photosynthesis is limited by the lack of water, and in cold and the Arctic regions it is limited by the lack of both liquid water and sunlight over a long period of the year. When moving away from hot deserts and cold and arctic climates, trophic structures TS 1 change into TS 2 and the latter into TS 3. This sequence (TS 1 → TS 2 → TS 3) appears in regions of the world with different climatic transitions, all sharing increases of energy available: the increase of precipitation when moving away from deserts and the increase of sunlight when moving away from the poles. The same pattern of transitions associated with increasing levels of precipitation and sunlight is observed worldwide from TS 3 to TS 6 trophic structures.

To examine the role of energy in explaining shifts in community trophic structures, we explore the relationship between trophic structures and NPP independently of climate. The predictive power of models including NPP alone is lower than models including all climate variables. However, NPP alone can reproduce the general division between trophic structures (Fig. 4f). Adding NPP to models including climate variables do not increase their predictive power, which again is unsurprising given that climate variables explain ca 97% of NPP variance. But when NPP is added alongside climate it becomes by far the single most important variable, suggesting that community trophic structures are strongly associated with the chemical energy available for consumers.

The emergence of community trophic structure seems strongly related to plant production rates, a feature that NPP seeks to characterize. However, that NPP alone does not explain all the variation in trophic structures explained by climate, suggests that additionally to plant production other aspects of climate exert controlling effects (e.g. its effects on vegetation types or via physiological effects influencing the identity of the extant species available for colonization within

regional pools). The observed trophic-structure-NPP patterns are consistent with the idea that average scores of NPP present in any location act as a surrogate for the amount of resources available for consumption by resident animal communities (Phillips et al. 2008, Pellissier et al. 2018), including humans (Krausmann et al. 2013, Moore et al. 2019), whereas excesses scores of NPP over the mean would represent resources becoming available for harvest by external migrant species (Thorup et al. 2017), a pattern apparently reproduced across geological period (Thorup et al. 2021).

As many other macroecological data explorations, our analysis has several potential caveats. Firstly, the number of species per trophic guild is only an estimation of the energy processed by each guild and may vary substantially across regions and sets of organisms. Secondly, we use grid cells with more than 10 000 km², because this is the resolution of the data, and they are treated as homogeneous. However, there are major local sources of within-cell environmental and biological variation, especially in heterogeneous environments, which could add biases and noise to the results. Naturally, the expectation is that, at global scale, such within-cell variation would be smoothed out giving rise to first order gradients of biogeographical significance. Thirdly, despite using a long list of bird and mammal species the faunal list of each cell is a subset of the total number within communities and their distributions were estimated based on global species range maps, which are interpolated based on sets of rules and expert judgement. Therefore, they likely include a large quantity of false positives. Fourthly, the characterization of the feeding habits that underlies the classification of guilds and then of trophic structures is an approximation given that, for most species, feeding preferences are based on a single study. For many species, there are not even studies available, and their trophic habits are inferred from taxonomy. For instance, some *Thamnophilidae* and *Tyrannidae* are classified as specialized insectivores, but some can also eat vertebrates or fruits. Finally, most communities have been altered by human activities and the different species of birds and mammals are unlikely to be affected in the same way causing biases that are difficult to qualify let alone quantify.

These uncertainties will undoubtedly encourage others to do more and better. Notwithstanding, the implications of our results are important for several reasons. Firstly, given the close correspondence between the trophic structures and climate, it is possible to infer the range of expected numbers of species per guild in any given region thus providing a convenient baseline against which to compare observed patterns (Gillson et al. 2011). Measurements of the multivariate distance between observed trophic structures and expected trophic structures given the climate, also allow analysis of the departures from this relationship and exploration of the potential causes. Modelled trophic structures against climate variables can also be projected into the future, thus enabling assessments of climate change effects on the biogeography of feeding relationships among species.

Secondly, the analysis of trophic structures can be modified to contribute to species distribution modelling efforts

(Peterson et al. 2011). Essentially, the procedure described in this study can be used to identify the potential set of interaction links among species; analogous to approaches leading to the identification of the backbone of biotic interaction networks or meta-webs (Morales-Castilla et al. 2015, Albouy et al. 2019). Taking into account the trophic interactions of the target species, one can build co-occurrence networks (Araújo et al. 2011) or joint species distribution models (Norberg et al. 2019), also known as community-based species distribution models (Ferrier and Guisan 2006, Baselga and Araújo 2009), that exclude coexisting species for which no critical interaction is expected. Such an approach is common for species distribution models that use a small subset of covariate species known to be critical interactors (Araújo and Luoto 2007, Heikkinen et al. 2007, Meier et al. 2010, Fordham et al. 2013), but it has not yet been performed in models examining large numbers of coexisting species simultaneously.

All in all, the emergence of general patterns in the functional properties of animal communities at broad scales is likely to spur several additional questions and analyses, and support the emergence of food-web biogeography as a new sub-discipline of biogeography focusing on the analysis of the geographical distributions of trophic relationships among organisms.

Acknowledgements – The authors thank Dominique Gravel, Carlos Ramos, José Maria Rey and Alejandro Rozenfeld for discussion of several of the theoretical ideas developed in this manuscript.

Funding – MM was originally funded through 1/SAESCTN/ALENT-07-0224-FEDER-001755 led by MBA. Both are now funded through the Spanish Ministry of Science, Innovation and Universities (PGC2018-099363-B-I00).

Author contributions

Manuel Mendoza: conceptualization (equal); writing (equal); methodology (lead). **Miguel B. Araujo:** conceptualization (equal); writing (equal).

Transparent Peer Review

The peer review history for this article is available at <<https://publons.com/publon/10.1111/ecog.06289>>.

Data availability statement

The data and code necessary to reproduce the results of this research are available from the Dryad Digital Repository, <<https://doi.org/10.5061/dryad.nk98sf7st>> (Mendoza and Araujo 2022).

Supporting information

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

References

- Albouy, C. et al. 2019. The marine fish food web is globally connected. – *Nat. Ecol. Evol.* 3: 1153–1161.
- Araújo, M. B. and Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. – *Global Ecol. Biogeogr.* 16: 743–753.
- Araújo, M. B. and Pearson, R. G. 2005. Equilibrium of species' distributions with climate. – *Ecography* 28: 693–695.
- Araújo, M. B. and Rozenfeld, A. 2013. The geographic scaling of biotic interactions. – *Ecography* 37: 406–415.
- Araújo, M. B. et al. 2005. Validation of species–climate impact models under climate change. – *Glob. Change Biol.* 11: 1504–1513.
- Araújo, M. B. et al. 2008. Quaternary climate changes explain diversity among reptiles and amphibians. – *Ecography* 31: 8–15.
- Araújo, M. B. et al. 2011. Using species co-occurrence networks to assess the impacts of climate change. – *Ecography* 34: 897–908.
- Araújo, M. B. et al. 2019. Standards for distribution models in biodiversity assessments. – *Sci. Adv.* 5: eaat4858.
- Baselga, A. and Araújo, M. B. 2009. Individualistic vs community modelling of species distributions under climate change. – *Ecography* 32: 55–65.
- Baselga, A. and Araújo, M. B. 2010. Do community-level models describe community variation effectively? – *J. Biogeogr.* 37: 1842–1850.
- Beisner, B. E. et al. 2003. Alternative stable states in ecology. – *Front. Ecol. Environ.* 1: 376–382.
- Bennett, D. E. et al. 2021. The evolution of the rangeland trusts network as a catalyst for community-based conservation in the American West. – *Conserv. Sci. Pract.* 3: 257.
- Bezdek, J. C. 1981. Pattern recognition with fuzzy objective function algorithms. – Plenum Press.
- Boettiger, C. and Batt, R. 2020. Bifurcation or state tipping: assessing transition type in a model trophic cascade. – *J. Math. Biol.* 80: 143–155.
- Boulesteix, A.-L. et al. 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. – *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 2: 493–507.
- Breiman, L. 2001. Random forests. – *Mach. Learn.* 45: 5–32.
- Callaway, R. M. et al. 2002. Positive interactions among alpine plants increase with stress. – *Nature* 417: 844–848.
- Carletta, J. 1996. Assessing agreement on classification tasks: the kappa statistic. – *Comput. Linguist.* 22: 249–254.
- Carpenter, S. R. and Pace, M. L. 2013. Dystrophy and eutrophy in lake ecosystems: implications of fluctuating inputs. – *Oikos* 78: 3–14.
- Cornell, H. V. and Harrison, S. P. 2014. What are species pools and when are they important? – *Annu. Rev. Ecol. Evol.* 45: 45–67.
- Coutron, P. and Lejeune, O. 2001. Periodic spotted patterns in semi-arid vegetation explained by a propagation-inhibition model. – *J. Ecol.* 89: 616–628.
- Cramer, W. et al. 1999. Comparing global models of terrestrial net primary productivity (NPP): overview and key results. – *Global Change Biol.* 5: 1–15.
- Currie, D. J. et al. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. – *Ecol. Lett.* 7: 1121–1134.
- Davis, A. J. et al. 1998. Making mistakes when predicting shifts in species range in response to global warming. – *Nature* 391: 783–786.
- De Angelis, D. et al. 2021. Environmental and anthropogenic correlates of seasonal migrations in the Dinaric-Pindos brown bear population. – *J. Zool.* 314: 58–71.
- Demencol, P. et al. 2000. Abrupt onset and termination of the African humid period: rapid climate responses to gradual insolation forcing. – *Quat. Sci. Rev.* 19: 347–361.
- Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – *Ecography* 30: 609–628.
- Dunn, J. C. 1973. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. – *J. Cybern.* 3: 32–57.
- Ferrier, S. and Guisan, A. 2006. Spatial modelling of biodiversity at the community level. – *J. Appl. Ecol.* 43: 393–404.
- Ficetola, G. F. et al. 2017. Global determinants of zoogeographical boundaries. – *Nat. Ecol. Evol.* 1: 89.
- Fordham, D. A. et al. 2013. Adapted conservation measures are required to save the Iberian lynx in a changing climate. – *Nat. Clim. Change* 3: 899–903.
- García, R. A. et al. 2014. Multiple dimensions of climate change and their implications for biodiversity. – *Science* 344: 1247579.
- García-Valdés, R. et al. 2013. Chasing a moving target: projecting climate change-induced shifts in non-equilibrium tree species distributions. – *J. Ecol.* 101: 441–453.
- Gleason, H. A. 1926. The individualistic concept of the plant association. – *Bull. Torrey Bot. Club* 53: 7.
- Gravel, D. et al. 2011. Trophic theory of island biogeography. – *Ecol. Lett.* 14: 1010–1016.
- Grubinger, T. et al. 2017. Package 'evtree.' – <<https://cran.r-project.org/web/packages/evtree/index.html>>.
- Hawkins, B. A. et al. 2003. Energy, water and broad-scale geographic patterns of species richness. – *Ecology* 84: 3105–3117.
- Heikkinen, R. K. et al. 2007. Biotic interactions improve prediction of boreal bird distributions at macro-scales. – *Global Ecol. Biogeogr.* 16: 754–763.
- Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – *Int. J. Climatol.* 25: 1965–1978.
- Holdridge, L. R. 1947. Determination of world plant formations from simple climatic data. – *Science* 105: 367–368.
- Holt, B. G. et al. 2013. An update of Wallace's zoogeographic regions of the world. – *Science* 339: 74–78.
- Hubell, S. P. 2005. Neutral theory in community ecology and the hypothesis of functional equivalence. – *Funct. Ecol.* 19: 166–172.
- Humboldt, A. von 1807. Ideen zu einer Geographie der Pflanzen nebst einem Naturgemälde der Tropenländer. – *Abt., Bd.*
- Imhoff, M. et al. 2004. Global patterns in human consumption of net primary production. – *Nature* 2: 1–16.
- IUCN 2014. The IUCN Red List of threatened species. Ver. 2014-3. – <www.iucnredlist.org>.
- Janis, C. M. 1990. Correlation of cranial and dental variables with dietary preferences in mammals: a comparison of macropodoids and ungulates. – *Mem. Queensl. Mus.* 28: 349–366.
- Konar, B. and Estes, J. A. 2003. The stability of boundary regions between kelp beds and deforested areas. – *Ecology* 84: 174–185.
- Krausmann, F. et al. 2013. Global human appropriation of net primary production doubled in the 20th century. – *Proc. Natl Acad. Sci. USA* 110: 10324–10329.

- Krętownski, M. and Grześ, M. 2005. Global learning of decision trees by an evolutionary algorithm. – In: Saeed, K. and Pejas, J. (eds), *Information processing and security systems*. Springer, pp. 401–410.
- Lejeune, O. et al. 2002. Localized vegetation patches: a self-organized response to resource scarcity. – *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.* 66: 10901.
- Liaw, A. and Wiener, M. 2002. Classification and regression by Random Forest. – *R News* 2: 18–22.
- Loreau, M. 2004. Does functional redundancy exist? – *Oikos* 104: 606–611.
- Loreau, M. 2020. The ecosystem: superorganism, or collection of individuals? – In: Dobson, A. et al. (eds), *Unsolved problems in ecology*. Princeton Univ. Press, pp. 218–224.
- Madon, B. et al. 2013. Community-level vs species-specific approaches to model selection. – *Ecography* 36: 1291–1298.
- Meier, E. S. et al. 2010. Biotic and abiotic variables show little redundancy in explaining tree species distributions. – *Ecography* 33: 1038–1048.
- Mendoza, M. and Araújo, M. B. 2019. Climate shapes mammal community trophic structures and humans simplify them. – *Nat. Commun.* 10: 5197.
- Mendoza, M. and Araújo, M. B. 2022. Data from: Biogeography of bird and mammal trophic structures. – Dryad Digital Repository, <<https://doi.org/10.5061/dryad.nk98sf7st>>.
- Mendoza, M. et al. 2004. Emergence of community structure in terrestrial mammal-dominated ecosystems. – *J. Theor. Biol.* 230: 203–214.
- Moore, C. et al. 2019. Human appropriated net primary productivity of complex mosaic landscapes. – *Front. For. Global Change* 2: 38.
- Moore, J. C. and De Ruiter, P. C. 2012. Energetic food webs: an analysis of real and model ecosystems. – Oxford Univ. Press.
- Morales-Castilla, I. et al. 2015. Inferring biotic interactions from proxies. – *Trends Ecol. Evol.* 30: 347–356.
- Morrone, J. J. 2015. Biogeographical regionalisation of the world: a reappraisal. – *Aust. Syst. Bot.* 28: 81.
- Norberg, A. et al. 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. – *Ecol. Monogr.* 89: 1–23.
- Parker, G. A. and Smith, J. M. 1990. Optimality theory in evolutionary biology. – *Nature* 348: 27–33.
- Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – *Global Ecol. Biogeogr.* 12: 361–371.
- Pellissier, V. et al. 2018. Niche packing and expansion account for species richness–productivity relationships in global bird assemblages. – *Global Ecol. Biogeogr.* 27: 604–615.
- Peterson, A. T. et al. 2011. Ecological niches and geographic distributions. – *Monographs in population biology*, no. 49. Princeton Univ. Press.
- Phillips, L. B. et al. 2008. Evaluating the species energy relationship with the newest measures of ecosystem energy: NDVI versus MODIS primary production. – *Remote Sens. Environ.* 112: 4381–4392.
- Post, E. S. 2013. Ecology of climate change: the importance of biotic interactions. – Princeton Univ. Press.
- Post, E. et al. 1999. Ecosystem consequences of wolf behavioral response to climate. – *Nature* 401: 905–907.
- Pyke, G. H. 1984. Optimal foraging theory: a critical review. – *Annu. Rev. Ecol. Syst.* 15: 523–575.
- Ricklefs, R. E. 2008. Disintegration of the ecological community. – *Am. Nat.* 172: 741–750.
- Rodrigues, P. et al. 2015. A biogeographical regionalization of Angolan mammals. – *Mamm. Rev.* 45: 103–116.
- Romdal, T. S. et al. 2013. Life on a tropical planet: niche conservatism and the global diversity gradient. – *Global Ecol. Biogeogr.* 22: 344–350.
- Scheffer, M. and Carpenter, S. R. 2003. Catastrophic regime shifts in ecosystems: linking theory to observation. – *Trends Ecol. Evol.* 18: 648–656.
- Sitch, S. et al. 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. – *Global Change Biol.* 9: 161–185.
- Soberón, J. and Nakamura, M. 2009. Niches and distributional areas: concepts, methods and assumptions. – *Proc. Natl Acad. Sci. USA* 106: 19644–19650.
- Solé, R. V. and Bascompte, J. 2012. Self-organization in complex ecosystems. – *Monographs in Population Biology*, Princeton Univ. Press.
- Storch, D. et al. 2018. The more-individuals hypothesis revisited: the role of community abundance in species richness regulation and the productivity–diversity relationship. – *Ecol. Lett.* 21: 920–937.
- Strobl, C. et al. 2008. Conditional variable importance for random forests. – *BMC Bioinformatics* 9: 307.
- Thorup, K. et al. 2017. Resource tracking within and across continents in long-distance bird migrants. – *Sci. Adv.* 3: e1601360.
- Thorup, K. et al. 2021. Response of an Afro-Palaearctic bird migrant to glaciation cycles. – *Proc. Natl Acad. Sci.* 118: e2023836118.
- Van Geest, G. J. et al. 2003. Vegetation abundance in lowland flood plan lakes determined by surface area, age and connectivity. – *Freshwater Biol.* 48: 440–454.
- von Humboldt, A. 2013. *Ideen zu einer Geographie der Pflanzen*. – Dogma, Germany.
- Whittaker, R. J. et al. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. – *J. Biogeogr.* 28: 453–470.
- Wilman, H. et al. 2014. EltonTraits 1.0: species-level foraging attributes of the world's birds and mammals. – *Ecology* 95: 2027–2027.
- Woodward, F. I. and Williams, B. G. 1987. Climate and plant distribution at global and local scales. – *Vegetatio* 69: 189–197.
- Zhang, J. 2009. Energy flows in complex ecological systems: a review. – *J. Syst. Sci. Complex.* 22: 345–359.