

Research article

A latitudinal signal in the relationship between species geographic range size and climatic niche area

Tad A. Dallas and Andrew Kramer

T. A. Dallas (<https://orcid.org/0000-0003-3328-9958>) ✉ (tad.a.dallas@gmail.com), Dept of Biological Sciences, Louisiana State Univ., Baton Rouge, LA, USA and Dept of Biological Sciences, Univ. of South Carolina, Columbia, SC, USA. – A. Kramer (<https://orcid.org/0000-0002-9031-6580>), Dept of Integrative Biology, Univ. of South Florida, Tampa, FL, USA.

Ecography

2022: e06349

doi: 10.1111/ecog.06349

Subject Editor: Timothy Keitt
Editor-in-Chief: Miguel Araújo
Accepted 31 August 2022



Species with broader niches may have the opportunity to occupy larger geographic areas, assuming no limitations on dispersal and a relatively homogeneous environmental space. Here, we use data on a large set of mammal ($n = 1225$), bird ($n = 1829$) and tree ($n = 341$) species to examine the 1) relationship between geographic range size and climatic niche area, 2) influence of species traits on species departures from this relationship and 3) sensitivity of these relationships to how species range size and climatic niche area are estimated. We find positive geographic range size–climatic niche area relationships for all taxa, with residual variation dependent on latitude, and differing from a null model for mammals and birds, but not for trees. Together, we provide support for this general macroecological relationship which is dependent on space, weakly influenced by species traits, and different enough from a null model to suggest that geographic and demographic processes are important.

Keywords: convex hull, geographic range size, latitudinal gradient, species range estimation

Introduction

Geographically widespread species tend to also have larger ecological niches relative to smaller-ranged species (Brown 1984, Slatyer et al. 2013, Yu et al. 2017). Niche space can be defined in terms of niche breadth (i.e. the range of resources that a species uses Rolando (1990)) or climatic niche area (i.e. the range of climatic conditions the species occurs in Dallas et al. (2017)). Previous investigations into the relationship between species geographic range size and niche size have defined the niche in several ways, including the use of habitat (e.g. number biomes occupied), diet (e.g. number of different food types used) and environmental tolerance (e.g. elevational range) measures (Morin and Lechowicz 2013, Morueta-Holme et al. 2013, Slatyer et al. 2013). The difficulties in defining and quantifying species geographic range (Gaston and Fuller 2009) and climatic niche (Rolando 1990, Violle and Jiang 2009, Fordyce et al. 2016, Machovsky-Capuska et al. 2016) size have contributed to the mixed support for these



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.

scaling relationships (Gaston and Spicer 2001, Morueta-Holme et al. 2013, Slatyer et al. 2013). However, two fairly recent studies found strong support across many different species for positive geographic range size–niche size relationships (Slatyer et al. 2013, Kambach et al. 2018), suggesting that these relationships may be quite general, but sensitive to spatial scale (Kambach et al. 2018) and spatial autocorrelation in climatic conditions (Moore et al. 2018).

Examining the relationship between geographic range size and climatic niche area also affords another interesting avenue; examining the effects of different geographic range size and niche area estimation procedures on subsequent scaling relationships. With respect to geographic range size estimation, Gaston and Fuller (2009) differentiated area of occupancy (AOO) from extent of occurrence (EOO), where AOO more closely relates to within range habitat utilization (e.g. fraction of sampled sites within a species range that are occupied), whereas measures of geographic range area more closely relate to EOO. A consensus on best practices for estimation of species geographic range size (Graham and Hijmans 2006, Gaston and Fuller 2009) and niche area (Blonder et al. 2014, Swanson et al. 2015) has yet to develop. However, given that distinct methods for estimating either geographic range size or climatic niche area have the same goal, it would be expected that different estimates of geographic range size would be correlated, as would estimates of niche area (but see Gaston and Fuller 2009). As a result, qualitatively consistent relationships between geographic range size and climatic niche area would likely be observed regardless of area measurement, but the degree of support for these relationships as a function of area estimation approach could still be quite variable. Therefore, careful justification of area estimation technique, or the use of multiple approaches, is important for ecological inference.

Apart from differences in measurement, the relationship between geographic range size and climatic niche area may be sensitive to another issue. Due to spatially autocorrelated environmental conditions, the scaling of species geographic range size and climatic niche area may simply be an artifact (Moore et al. 2018). That is, more widespread species are likely to encounter a larger range of climatic conditions compared to species with more restricted geographic ranges (Saupe et al. 2019). This is normally viewed from the niche-perspective, in that species with larger climatic niches will be capable of colonizing a larger set of geographic locations (Pulliam 2000). Disentangling the underlying relationship from an artifact of a spatial sampling process is difficult with observational data, as the niche and the geographic distribution are inherently linked through the observation process (Pulliam 2000, Colwell and Rangel 2009). Mathematical model simulations can start to address this issue, as can comparing observations in natural systems to a representative null model.

From a practical perspective, the potential uncoupling of geographic range size and climatic niche area as a function of spatial processes can lead to species diverging from the expected global relationship, which would consider data on all

available species within some taxonomic group(s) to estimate the relationship between geographic range size and climatic niche area. However, the availability of colonizable land and climatic niche area is not uniformly distributed across space (discussed for plants in Sheth et al. 2020), suggesting that spatial gradients can constrain geographic range size–climatic niche area relationships even when all species are functionally neutral (Blackburn and Gaston 1997). The structure of the land masses also matters, as narrow bottlenecks present clear dispersal barriers for some species (Brown and Maurer 1989). This creates an interesting possibility; the availability of land area and climatic niche space in a given area places fundamental constraints on the resulting geographic range size–climatic niche area relationship (Baselga et al. 2012). This could suggest the existence of a spatial signal in divergence from the global geographic range size–climatic niche area relationship, potentially driven by geographic areas of high discordance (e.g. large geographic space with low climatic heterogeneity).

But while a geographic range size–climatic niche area relationship might vary across spatial gradients and be expected under neutral assumptions, divergence from the expected scaling relationship across a wide number of species might point to an underlying species attribute associated with divergence (e.g. dispersal ability (Pagel et al. 2020)). Relating the residual deviations from the expected relationship to geographic covariates and species traits may provide insight into when and where relationships between species geographic range size and climatic niche area are weakened. For instance, such a process could identify species with large geographic ranges and smaller than expected niche areas. This would tend to occur in geographic locations which remain climatically similar across large geographic spaces. That is, a species may specialize on a very common set of environments, leading to a rather small climatic niche area and a large geographic range size. This specialization may relate to species traits, especially traits that are associated with geographic range size such as body size in mammals (Gaston and Blackburn 1996) and height in plants (Mashau et al. 2021).

Here we examine the relationship between geographic range size and climatic niche area using three large datasets on mammal ($n = 1225$), bird ($n = 1829$) and tree ($n = 341$) species distributed across the Americas. The large number of species and variety of taxa provide the chance to examine the generality this relationship. By spanning large latitudinal ranges, these data can also address the dependence of the relationship on geographic and climatic constraints. A null model is used to determine whether variation around the general relationship exceeds what might be expected due to change and latitudinal variation in land area and niche space. We will also consider the influence of species traits on the geographic range size–climatic niche area relationship and assess the sensitivity to different range size or climatic niche area estimation approaches. By combining these analyses on an extensive dataset we are able to extend understanding of the generality of the geographic range size–climatic niche area relationships and address the role of potential fundamental constraints on the relationship.

Methods

Data sources

We obtained species occurrence and trait data for mammal, tree and bird species from freely available data sources (described below). To include as many species as possible, we queried species occurrence records from the Global Biodiversity Information Facility – a species occurrence database – using the R package ‘rgbif’ (Chamberlain et al. 2016) for all mammal species listed in PanTHERIA (Jones et al. 2009), a mammal trait database. Species occurrence records were checked for quality using the package ‘scrubr’, which removed occurrences with missing, uncertain or unlikely (0°N 0°W) latitude and longitude values (Chamberlain 2016). We obtained bird occurrence data from eBird (Sullivan et al. 2009), and bird trait data from Myhrvold et al. (2015). Lastly, tree occurrence data was obtained from the US Department of Agriculture Forest Inventory and Analysis database (Bechtold et al. 2005), and tree trait data was obtained from TRY (Kattge et al. 2011). It is important that these data sources fundamentally differ in several key aspects. First, the FIA tree data are limited to the continental United States, meaning that range estimation and climatic niche area may be underestimated for species that extend beyond those borders. Second, some data are from opportunistic sampling, meaning that sampling bias across space or for certain species may be present, while other data sources (e.g. FIA tree data) come from systematic sampling efforts. This is a strength of using multiple data sources with different idiosyncracies, as the ideal data do not exist, but multiple lines of support from different data sources for a single macroecological relationship can provide consensus or highlight important differences among data types or taxa. See the Supporting information for further discussion and maps of species occurrences.

Relevant species traits are described below (and also in Table 1), but largely center on life history traits influencing the range of habitats a species can occupy or the pace of

life (e.g. birth rate) of a given species, taken as a subset from available trait data sources. Trait data for some species were unavailable, resulting in variable trait data coverage (e.g. 90% coverage on mammal body sizes, but only around 30% coverage on mammal lifespan). However, our modeling approach still allowed us to use cases of missing data.

Through this approach, we obtained data on 1277 mammal species, 1885 bird species and 352 tree species, resulting in nearly 4.3 million species occurrence records. Species occurrence records were filtered to only those occurrences in the Americas, in order to avoid complications in estimating geographic range size across large amounts of inhospitable habitat (e.g. ocean). Further, species with fewer than four unique geographic occurrence records or fewer than four unique climatic niche values were not considered in the analyses, slightly reducing the final number of mammal ($n=1225$), bird ($n=1829$) and tree ($n=341$) species included in the analyses.

Estimation of geographic range and climatic niche area

Many methods have been developed to estimate species geographic range and climatic niche area (Quinn et al. 1996, Burgman and Fox 2003, Lichti and Swihart 2011), each of which makes tacit assumptions about the structure of the climatic niche or the spatial distribution of a species across a landscape (Fig. 1). For instance, the convex hull approach may potentially overestimate geographic range area as a result of the limitation that the minimum bounding polygon of the species range can only contain convex angles, leading to geographic areas which may contain regions of geographic space where a species cannot exist. More restrictive approaches – such as alpha hulls – attempt to account for this, but, in doing so, may (Darroch and Saupe 2018) or may not (Burgman and Fox 2003) be more prone to issues with sampling or detection bias. There is presently no clear optimal method for estimating species geographic range sizes

Table 1. Species traits examined for their associations with residuals of the relationship between geographic range size and climatic niche area.

Species group	Trait	Definition	Units
Mammals	Body mass	Mass of adult host	$\log(1 + g)$
	Diet breadth	Number of dietary categories eaten by host species	#
	Habitat breadth	Number of habitats occupied (ground dwelling, aquatic, fossorial)	#
	Home range size	Average area of habitat utilized for host species	Km ²
	Litter size	Number of offspring per litter	#
	Maximum age	Maximum age for host species	months
	Population density	Number individuals per square km	#/km ²
	Trophic level	Herbivore, carnivore or omnivore	factor
	Trees	Dispersal distance	Average dispersal distance
Plant height		Average height of plant	m
Seed size		Average mass of plant seed	mg
Birds	Body mass	Average adult body mass	$\log(1 + g)$
	Clutch size	Number of offspring per clutch	#
	Egg mass	Mass of egg	g
	Incubation period	Length of incubation period	days
	Lifespan	Maximum age for bird species	years

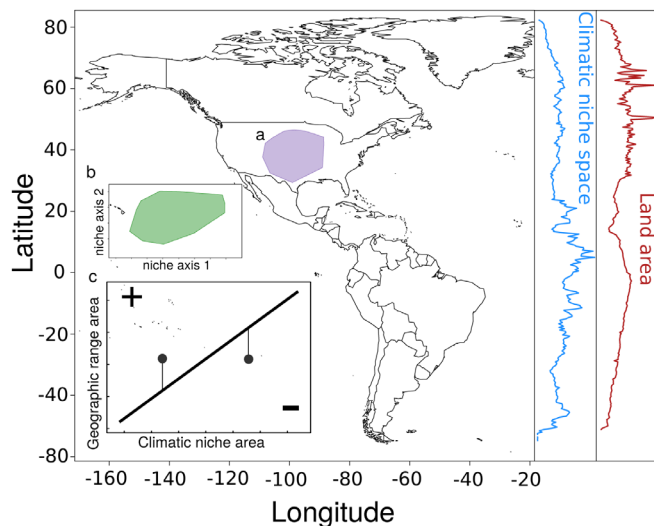


Figure 1. Residual variation from the relationship between a species geographic range size (spatial polygon in panel a) and corresponding climatic niche area (polygon in panel b) may be associated with species traits or spatial structure. These residuals (depicted in panel c) represent situations where geographic range area is larger than expected given the climatic niche size (indicated with a red dot and a '+' symbol) or where geographic area is smaller than expected given the climatic niche area (indicated with a blue dot and a '-' symbol). However, a spatial signal in the relationship between geographic range size and climatic niche area may exist, as latitudinal variation exists in both available climatic niche space (blue line) and geographic land area (red line).

or climatic niche areas. As such, we use three well-established methods; convex hull, alpha hull (with two different alpha parameterizations) using the 'alphahull' R package (Pateiro-Lopez et al. 2016), and standard ellipse area using the 'siar' R package (Parnell and Jackson 2013). In the main text, we estimate species ranges using the convex hull, as it is well-established and does not require parameterization like other methods (e.g. α parameter of alpha hulls). In the Supporting information, we discuss the sensitivity of the scaling relationship between geographic range size and climatic niche area to the measure used to estimate geographic and climatic niche areas. We find similar relationships regardless of range estimation approach, though the shape of the relationship and resulting residuals (and models fit to residuals) differed, highlighting the sensitivity of macroecological relationships to statistical approach, while also providing support for the generality of the positive relationship between geographic range size and climatic niche area. Geographic range area was calculated as the minimum convex hull of the species occurrences accurately accounting for the spherical shape of the globe using the *areaPolygon* function in package 'geosphere' (Hijmans 2021).

We operationalize the species niche as the set of climatic space a species occupies (most akin to a realized niche (Soberon 2007)). Climatic niche area was determined by first translating the multivariate climate space into a two-dimensional space comparable to geographic space. To do

this, we accessed the set of 56 BioClim/WorldClim variables (Hijmans et al. 2005) at 2.5 arc-degree resolution through the 'raster' R package. We then calculated the first two principal components (PCA), explaining over 77% of the total global climatic variation in a two-dimensional space (Kriticos et al. 2014, Kambach et al. 2018). The WorldClim variables ($n = 36$), containing monthly information on minimum and maximum temperature and precipitation, and the BioClim variables ($n = 19$), containing derived quantities such as temperature seasonality and mean annual precipitation, are well-tested and represent the best available large scale climatic data (Barbet-Massin and Jetz 2014). In terms of describing the species niche, this approach reduces the global climate to a small number of dimensions, and may not capture species-specific niche requirements. For instance, hibernating species may respond more strongly to climatic conditions in spring instead of in winter. However, by compressing monthly temperature and precipitation data down to two axes we have explained the majority of relevant climatic variation – at least for large-scale patterns across species entire geographic ranges. The first climatic PCA axis largely corresponds to temperature covariates, while the second PCA axis corresponds more to precipitation covariates (Supporting information).

Latitudinal variation in potential geographic range size–climatic niche area relationship

There is a clear latitudinal gradient in the amount of available land area and climatic niche space across the Americas (Fig. 1). We examined this variation by dividing the Americas into latitudinal bands (0.045 degree resolution), and calculating the total geographic land area and climatic niche area available if a species were to occupy that entire band. This creates a series of related points in the phase space of geographic range size and climatic niche area, which corresponds to the strongest possible relationship that could be observed if a species only occupied a given latitudinal band.

The importance of longitudinal variation and range aspect

The focus here on latitudinal variation may belie the importance of longitudinal gradients in geographic and climatic space. We start to address this by calculating the total available land area across latitudinal bands, but it is possible that the geographic range size–climatic niche area relationship – and subsequent departures from this relationship – could be a signal of the combination of latitudinal and longitudinal variation in species ranges. To explore this, we calculated the *range aspect*, quantified as the latitudinal range of a species divided by the longitudinal range of the species. *Range aspect* was significantly related to species latitudinal range center, positively for mammals and birds, and negatively for tree species (Supporting information). We explore range aspect as a potential predictor of departures from geographic range size–climatic niche area relationships in the Supporting

information, where we do not find strong evidence for the inclusion of this predictor. However, it is important to recognize the importance of species geographic range shape and the degree of spatial autocorrelation in climatic variables in order to understand the mapping between geographic distribution and climatic niche (see Baselga et al. (2012) for a great example).

Geographic and species traits associated with residuals

We related square-root-transformed species geographic range size to square-root-transformed species climatic niche area using linear regression (Supporting information). The residuals from these simple linear models represent the divergence of each species from the overall relationship between geographic range size and climatic niche area. If this relationship produces residuals that are no different from a null model, this would suggest that the residual variation is more a function of spatial autocorrelation, sampling and measurement error or differential species distributions independent of species identity.

To examine the null distribution of residuals, we simulated species geographic distributions and climatic niches by sampling the empirical occurrence values across ranges of latitude, total latitudinal range and occurrence number that were comparable to the empirical data (Supporting information). That is, we selected a random latitudinal minimum (Uniform (−57, 57) for mammals and birds, Uniform (18, 48) for trees to match their empirical distributions), a random latitudinal range size (Uniform (1, 45)) and a random number of occurrence points (Uniform (5, 500)). Given these ranges, we assembled species distributions by sampling the empirical occurrence data for mammals, trees and birds separately. That is, we essentially created null species by randomly sampling species occurrence data bounded between a given latitudinal range, and then calculated species geographic range size and climatic niche breadth for the null species. For each species present in the empirical data, we generated a null species which sampled occurrence data for a given taxa bounded by latitudinal constraints and the number of species occurrence records. By comparing the distribution of residual values from both empirical and null geographic range size–climatic niche area relationships, we explore to what extent this macroecological pattern is simply an emergent property of spatially autocorrelated environmental data, or if geographic or trait variation can drive departures from the expected relationship. We compared null and empirical distributions of residual values using Kolmogorov–Smirnov tests.

We then related these residuals to geographic variables and species traits using a regression tree approach. To examine spatial structure of residual variation in geographic range size–climatic niche area relationships, we included the latitudinal centre of a species range as a covariate. With respect to species traits, we selected life history traits (e.g. body size, habitat breadth, trophic level, dispersal distance) that have been previously found to be related to species geographic

range size, climatic niche breadth or competitive ability (Table 1). Further, we also selected some traits related to species 'pace of life' or demography, including litter size, lifespan and seed mass (Table 1). While we are inherently limited by the traits for which we have suitable data coverage, we might expect species traits associated with geographic range size (e.g. body size) to lead to divergence away from the relationship between geographic range size and climatic niche area.

Gradient boosted machines, also known as boosted regression trees, are a flexible regression technique in which many weak learning decision trees are iteratively created, where each tree attempts to explain variation left over from the previous tree (Friedman 2002, Elith et al. 2008). When these trees are combined, they are able to handle collinear data, handle missing values and account for variable interactions (Friedman 2002, De'ath 2007, Elith et al. 2008). Models were trained in R using the 'gbm' package (Greenwell et al. 2022), and were internally five-fold cross validated on 80% of the data to avoid overfitting, while the remaining 20% test data was used to evaluate model performance. This cross-validation was performed for each of 10 trained models per species group, each on a different random subset of 80% of the data, in order to examine variation in model performance as a function of the sampled data. Model performance was assessed using Spearman's rank correlation coefficients between residuals from the relationship between geographic range size and climatic niche area that were predicted from the boosted regression tree model compared to the actual values observed in the test data.

The relative importance of species trait covariates was determined by permuting each predictor variable individually and measuring the associated reduction in model performance (Breiman 2001), with values scaled between 0 and 100. This produces a relative importance measure whose values all sum to 100, with larger relative contribution values corresponding to greater importance to model performance. The directionality of the effects of the top species trait covariates was visualized using partial dependence plots, which show the relative effect of each variable at the average values of the other covariates (Elith et al. 2008).

Results

The relationship between geographic range and climatic niche area

Significantly positive relationships were observed between geographic range area and climatic niche area (Fig. 2) for mammals ($\beta=4.21$, $p < 0.0001$, adjusted $R^2=0.44$), trees ($\beta=2.77$, $p < 0.0001$, adjusted $R^2=0.41$) and birds ($\beta=4.72$, $p < 0.0001$, adjusted $R^2=0.60$). This supports previous findings suggesting the generality of this relationship (Slatyer et al. 2013, Kambach et al. 2018). Residuals from this linear relationship were detectably different in their distribution from a null expectation generated from sampling empirical occurrence data for randomly selected latitudinal bands

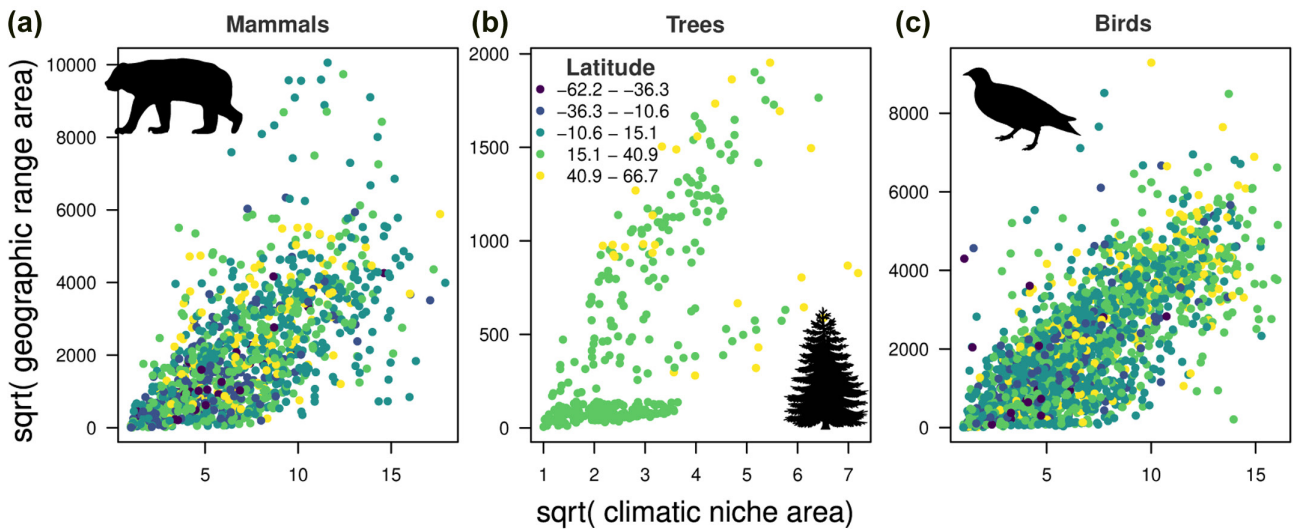


Figure 2. The relationship between a species geographic range size (square-root-transformed) and climatic niche area (square-root-transformed) for a set of 3395 species of (a) mammals, (b) trees and (c) birds. Point color is based on species' latitudinal centroids.

and number of occurrence points for mammals ($D=0.147$, $p < 0.0001$) and birds ($D=0.058$, $p=0.004$), but not significantly different for tree species ($D=0.090$, $p=0.126$) (Fig. 3), suggesting that the null distributions were incapable of capturing the true residual distributions for mammals and birds. The observed positive relationship between geographic range size and climatic niche area was maintained when other methods were used to estimate species geographic and climatic niche area as well, though comparisons between different area estimation methods for geographic range (Supporting information) or climatic niche area (Supporting information) resulted in more weakly related estimates (Supporting information). Across a latitudinal gradient in the Americas, the potential relationship space between geographic range size and climatic niche area varied drastically (Fig. 4), suggesting

that the slope of the relationship may be determined – at least in part – by the latitudinal range where the species is found.

Geographic and species traits associated with residuals

Residuals from the relationship between species geographic range size and climatic niche area were fairly well-predicted by species traits (Table 1) for mammal ($\bar{\rho} = 0.53 \pm 0.07$), tree ($\bar{\rho} = 0.80 \pm 0.02$) and bird ($\bar{\rho} = 0.55 \pm 0.03$) species (Fig. 5). Latitudinal centre was the most important covariate in boosted regression tree models for all species groups, suggesting a strong latitudinal signal on the geographic range size–climatic niche area relationship. This held true when the correlated spatial predictor of range aspect was included,

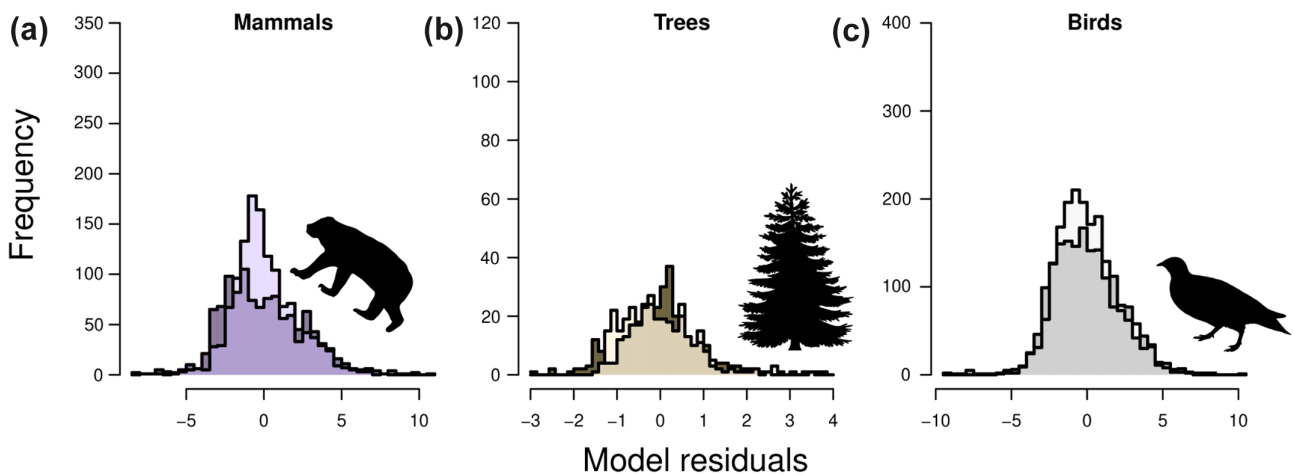


Figure 3. Empirical (darker colors) and null (lighter colors) distributions of residuals for (a) mammal, (b) tree and (c) bird species. Mammal and bird residual distributions were significantly different from the null expectation, while the tree residual distribution was not. However, all empirical and null distributions appear quite similar to one another. The null model was run 5000 times for each species group, here we sample the nulls randomly to match the number of species in each group for easier comparison.

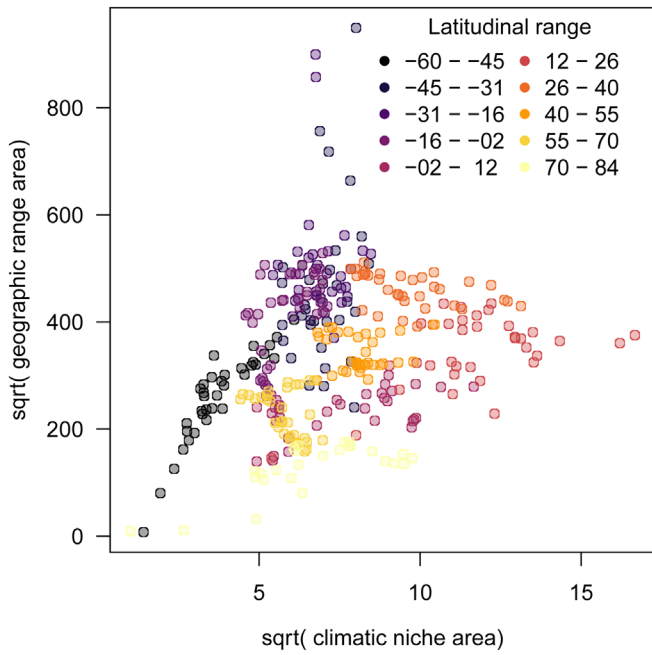


Figure 4. The potential relationship between geographic range size and climatic niche area is constrained by the amount of available land to colonize and the amount of available niche space. Here, each point is a latitudinal band, corresponding to the amount of land area and climatic niche area in that band. This demonstrates a clear latitudinal pattern in the relationship between these two variables, suggesting an underlying cause of the observed latitudinal signal.

except in the case of residuals for tree species which were also highly influenced by range aspect (Supporting information). For tree and bird species, the importance of latitudinal range centre dominated the model performance, with the remaining species traits contributing little to model performance (Fig. 5). However, for mammal species, species body mass was important for model performance (Fig. 5), potentially through the covariance between mammalian body size and latitude (e.g. Bergmann’s rule; (Ashton et al. 2000)). This effect was clear from the partial dependence plots, which examine the relative effect of each covariate on the residual variation in the geographic range size–climatic niche area relationship (Fig. 6). In these partial dependence plots, a positive relationship with a predictor would indicate that the predictor is associated with the situation where species have larger geographic ranges given their climatic niche area relative to the fit model with all species, with negative relationships indicating the opposite (i.e. species diverge from the relationship have a smaller geographic range than expected).

The importance of latitudinal centre to residual variation in geographic range size–climatic niche area relationships was not a result of latitudinal variation in geographic range size or climatic niche area, as there was no clear relationship between latitudinal centre and either geographic range size or climatic niche area (Supporting information). This suggests that the relationship between climatic niche area and geographic range size is influenced by latitude, but both area estimates are not strongly related to species latitudinal centre independently. Dividing the Americas into latitudinal bands, it becomes apparent that there is a latitudinal signal in the potential geographic range size–climatic niche area relationships that could emerge for a

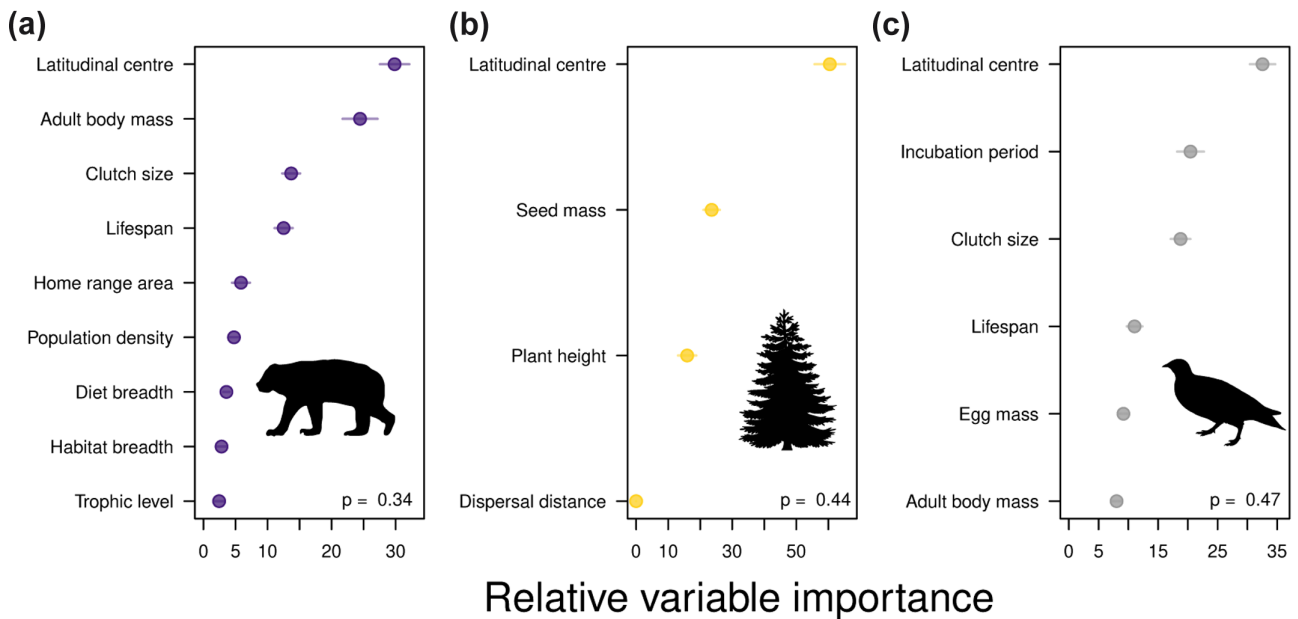


Figure 5. Relative importance values (mean and standard deviation) obtained from boosted regression tree models, which relate residual variation from geographic range size–climatic niche area relationships for mammals, trees and birds to species traits and latitudinal centre. The importance of species latitudinal range position suggests a strong signal of spatial processes.

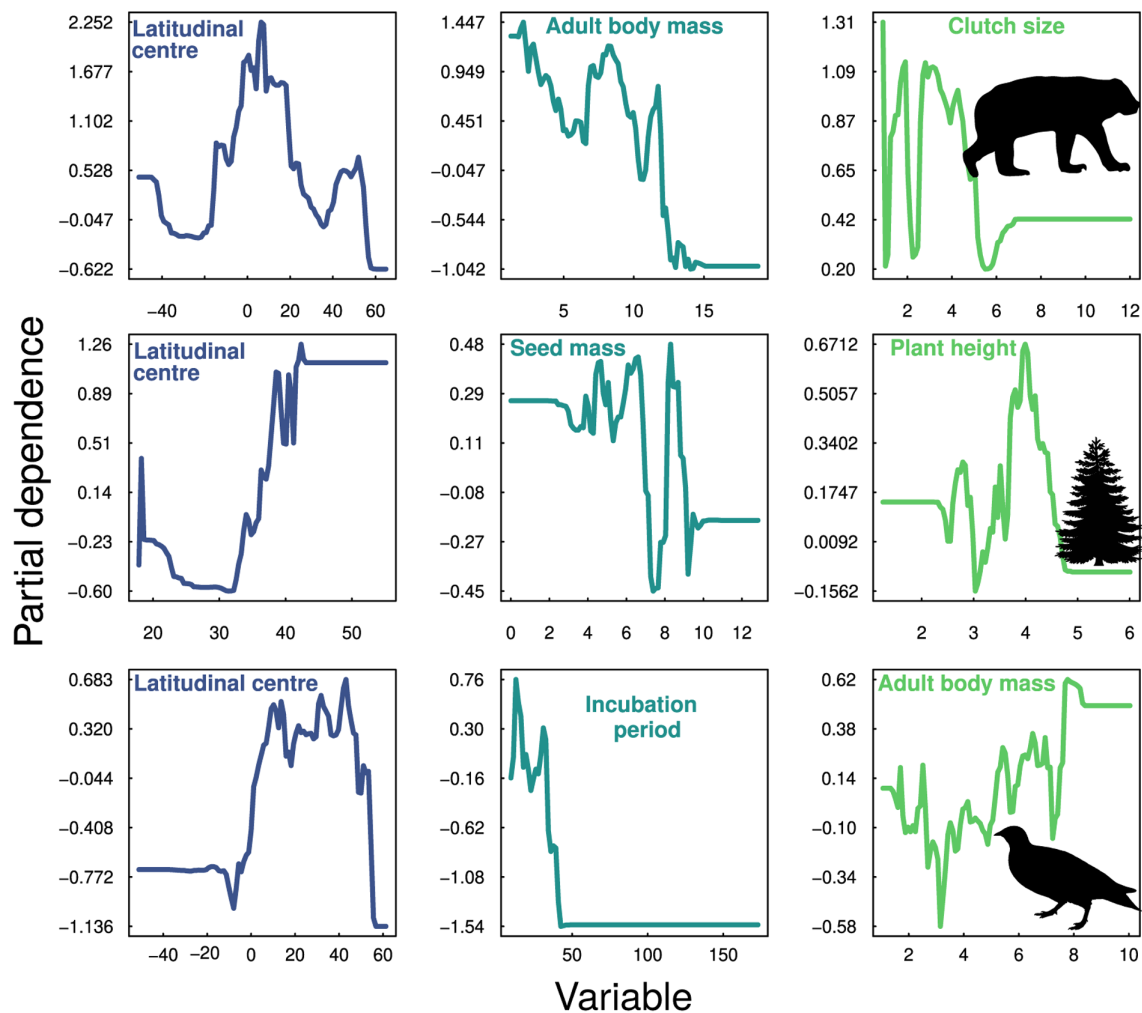


Figure 6. Partial dependence plots for the top three predictors in boosted regression tree models for mammals (top panel), trees (middle panel) and birds (bottom panel). Each panel represents the effect of the highlighted variable on the resulting residual variation in the geographic range size–climatic niche area relationship, where positive (negative) values correspond to a geographic range size which is smaller (larger) than expected given the estimated climatic niche area.

given species with a contiguous range (Fig. 4). Specifically, there is a positive relationship between geographic range area and climatic niche area across latitudinal bands, and the latitudinal range considered inherently influences the potential relationship between geographic range area and climatic niche area (e.g. compare latitudinal ranges 70–84 and –60 to –45 in Fig. 4). This does not inherently suggest that the residual variation in the geographic range size–climatic niche area relationship will be latitudinally structured, but simply that the distribution of potential climatic niche space and available land area does contain a latitudinal signal (Fig. 1 and 4).

A positive effect of latitudinal range centre suggests that species distributed at higher absolute latitudes tend to have larger geographic ranges than expected given their climatic niche sizes. However, given that the modeling approach used accounts for variable interactions and non-linear relationships, we see clear non-linear relationships between latitudinal range centre and residuals from the geographic range

size–climatic niche area relationship, with clear differences among taxa as well (Fig. 6). This is especially pronounced for mammals, as latitudinal centre is important to model fit, but the partial dependence plot highlights the non-linearity in the residuals (Fig. 6). A negative relationship between mammal body mass and the residual variation in the geographic range size–climatic niche area relationship suggests that larger-bodied mammals tend to have larger climatic niches than expected given their geographic range size (Fig. 6). This places latitudinal range centre and body size showing contrasting relationships with the residuals, despite the generally positive relationship between body mass and latitude. However, the relationship with body mass was quite variable, as we observed clear differences in the effect of body mass on residual variation of the geographic range size–climatic niche area relationship between mammals (negative relationship) and birds (positive relationship) (Fig. 6). These relationships were further supported when estimating geographic range size and climatic niche area using standard ellipse areas

(SIAR), though the use of alpha hulls resulted in reduced model performance (Supporting information).

Discussion

Geographic range size was positively related to climatic niche area for all species groups, supporting previous studies claiming this is a consistent ecological relationship (reviewed in Slatyer et al. 2013). The scaling between geographic range size and climatic niche area may stem from the close relationship between a species niche and its corresponding distribution (Pulliam 2000), though this tacitly assumes that species with broader climatic niches will inherently have larger geographic range areas, owing in part to the commonly observed spatial autocorrelation in climatic conditions (Moore et al. 2018). The assumed relationship between geographic range size and climatic niche area was supported by both our analyses and the null model, which generated residual variation which largely matched the empirical data. However, empirical residual distributions did differ from the null distributions for mammals and birds, where null distributions tended to have lower variance around the mean. We were also able to detect a latitudinal signal in the residual variation of the geographic range size–climatic niche area relationship. While this spatial structure appeared to take different forms for the 3 groups of species, the importance of latitude was likely a function of the distribution of available land area and climatic niche space, caused by the latitudinal decoupling between available land area and climatic niche space. Together, our findings suggest that geographic range size positively scales with climatic niche area, and that the residual variation in this relationship may be explained by species ecology and geography.

The existence of latitudinal structure in species deviations from a general geographic range size–climatic niche area relationship is an important finding, indicating a clear geographic constraint that outweighs a suite of species traits. The existence of spatial structure in available geographic range size and climatic niche area is an important consideration as well (Fig. 4), as available land area and niche area could fundamentally constrain the geographic range size–climatic niche relationship along spatial gradients. Further, there is mixed evidence suggesting that geographic range size (Stevens 1989) and climatic niche breadth (Vázquez and Stevens 2004) should scale positively with latitude. We did not find strong evidence for latitudinal scaling in square-root-transformed geographic range size or climatic niche area, where latitude was defined using the latitudinal centre for each species (Supporting information). However, we did see a clear increase in available geographic space in northern latitudes, while available niche space is highest in the tropics (Fig. 1). This creates a situation where the available geographic area and climatic niche space are latitudinally structured, constraining the possible relationship between geographic range size and climatic niche area for a given species occupying a contiguous geographic range (Fig. 4). Other methods of range and niche estimation may partially remove

this latitudinal structure (e.g. alpha hulls which allow discontinuous ranges and niches), though latitudinal centre tended to still be important when using alpha hulls here (Supporting information).

Apart from the spatial distribution of available land area and climatic niche space, many other relevant factors follow a latitudinal gradient. For instance, species diversity (Hillebrand 2004) and ecological interactions (Roslin et al. 2017) can be latitudinally-structured, and the latitudinally-structured residual variation in geographic range size–climatic niche area relationships may be a function of competition, natural enemies or dispersal limitation (Pagel et al. 2020). Further, geological and evolutionary processes underlying the distribution of species diversity and the shape of continents, may further influence the resulting variation in species geographic range size–climatic niche area relationships (Baselga et al. 2012). For example, the shape of continents appears to influence range aspect in the species considered here. Understanding the relative roles of species interactions, species traits and the entirely neutral constraints of available land area and climatic niche space is an important next step.

Due to spatially autocorrelated climatic gradients, increasing species geographic range is also likely to increase the climatic niche area. The spatial structuring of niche space is important when we consider that the global distribution of land area (and niche area) is non-random. This could lead to reduced niche breadth in the tropics solely as a result of the spatial structure of the available climatic niche space (Saupe et al. 2019). However, by focusing on explaining the residual variation in the relationship between species distribution and niche, we aimed to address the traits and spatial distributions of species which diverged from the expected scaling between geographic range size and climatic niche area. In our analyses, the same data used to estimate species geographic range data were used to identify the species climatic niche, an incredibly common practice in ecological niche modeling and macroecological analyses. While the use of independent data to estimate niche and geographic distribution would allow the separation of species niche and distribution, data availability constraints and the difficulty in translating laboratory-defined niche limits to natural systems are far greater issues than the perceived circularity in using spatial occurrence data to define both species geographic distribution and climatic niche. A final concern is that spatial biases in terms of sampling and detection could influence our results, especially if these biases contained a spatial signal. However, the use of data from both occurrence databases (e.g. Global Biodiversity Information Facility) and long-term, repeated, systematic surveys (e.g. USDA Forest Inventory and Analysis) suggests that differences in data 'quality' did not strongly influence the relationship between geographic range size and climatic niche area. Further, the important effect of species latitudinal centre across species groups and different data sources provides support for the generality of the relationship between geographic range size and climatic niche area.

Conclusions

Together, our findings provide further support for the strong relationship between geographic range size and climatic niche area (Slatyer et al. 2013), and suggest the existence of a clear, though idiosyncratic, spatial signal in this relationship for a diverse set of mammal, tree and bird species. This spatial signal is independent of latitudinal scaling relationships in geographic range size and climatic niche area, as we failed to detect a strong effect of latitude on either covariate. We posit that latitudinal structure in the availability of land area and climatic niche space is the underlying cause of the latitudinal structure in residual variation in the scaling relationships of geographic range size and climatic niche area for the mammals, trees and birds we examined. However, we also recognize that other latitudinally structured processes, such as competition, prey availability, natural enemies and dispersal limitation, may also contribute to species resulting geographic range size–climatic niche area relationships. Identifying species deviating from the general geographic range size–climatic niche area relationship may be important targets given climate change. For instance, species with larger niche areas than expected given their geographic distributions could be candidates for potential range expansion barring dispersal limitation. On the other hand, species with smaller niche areas than expected given their geographic distributions are likely occupying a common environmental space, but may be sensitive to loss or geographic shifts in the distribution of the narrow environmental niche space. For instance, the importance of mammalian body size to residual variation in the geographic range size–climatic niche area relationship could relate to latitudinal variation in body size (Bergmann's rule; (Ashton et al. 2000)) or be a signal of larger-bodied mammal distributions changing with human activity, being pushed from more moderate habitats to colder and drier climates (Pineda-Munoz et al. 2021), which may lead to a disconnect between the geographic range defined by a convex hull and the climatic niche. Monitoring species deviations from the overall relationship between geographic range size and climatic niche area may provide a way to identify species sensitive to range expansion or loss, and contribute to a trait-based understanding of geographic range size–climatic niche area scaling relationships.

Acknowledgements – We sincerely thank all individuals associated with the collection and curation of the open data on species occurrences (FIA, GBIF, eBird) and traits (Pantheria and the Amniote database) used in this manuscript. Discussions with Laura Antao greatly improved the manuscript. The study has been supported by the TRY initiative on plant traits (<www.trydb.org>). Silhouette images used in figures are from Phylopic (<<http://phylopic.org/>>).

Funding – This work has been performed with funding to Tad Dallas from the National Science Foundation (NSF-DEB-2017826) MacroSystems Biology and NEON-Enabled Science program.

Conflict of interest – The authors have no conflicts of interest to declare.

Author contributions

Tad A. Dallas: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Methodology (lead); Project administration (equal); Resources (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Andrew Kramer:** Conceptualization (supporting); Formal analysis (supporting); Investigation (supporting); Methodology (supporting); Resources (supporting); Software (supporting); Visualization (supporting); Writing – original draft (equal); Writing – review and editing (equal).

Transparent peer review

The peer review history for this article is available at <<https://publons.com/publon/10.1111/ecog.06349>>.

Data availability statement

Data and R code to reproduce the results are available from Figshare, <<https://doi.org/10.6084/m9.figshare.7964666>>. Data are available for eBird data (Sullivan et al. 2009), Forest Inventory and Analysis data (Woudenberg et al. 2010) (<www.fia.fs.fed.us/>) and the Global Biodiversity Information Facility data (Jones et al. 2009, Chamberlain et al. 2016). Authors should cite the original data sources.

Supporting information

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

References

- Ashton, K. G. et al. 2000. Is bergmanns rule valid for mammals? – *Am. Nat.* 156: 390–415.
- Barbet-Massin, M. and Jetz, W. 2014. A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. – *Divers. Distrib.* 20: 1285–1295.
- Baselga, A. et al. 2012. Global patterns in the shape of species geographical ranges reveal range determinants. – *J. Biogeogr.* 39: 760–771.
- Bechtold, W. A. et al. 2005. The enhanced forest inventory and analysis program-national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. – US Dept of Agriculture, Forest Service, Southern Research Station 85, p. 80.
- Blackburn, T. M. and Gaston, K. J. 1997. The relationship between geographic area and the latitudinal gradient in species richness in new world birds. – *Evol. Ecol.* 11: 195–204.
- Blonder, B. et al. 2014. The n-dimensional hypervolume. – *Global Ecol. Biogeogr.* 23: 595–609.
- Breiman, L. 2001. Random forests. – *Mach. Learn.* 45: 5–32.
- Brown, J. H. 1984. On the relationship between abundance and distribution of species. – *Am. Nat.* 124: 255–279.
- Brown, J. H. and Maurer, B. A. 1989. Macroecology: the division of food and space among species on continents. – *Science* 243: 1145–1150.

- Burgman, M. A. and Fox, J. C. 2003. Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. – *Anim. Conserv.* 6: 19–28.
- Chamberlain, S. 2016. *scrubr*: clean biological occurrence records. – <<https://cran.r-project.org/src/contrib/Archive/scrubr/>>.
- Chamberlain, S. et al. 2022. *rgbif*: Interface to the Global Biodiversity Information Facility API. R package ver. 3.7.3. – <<https://CRAN.R-project.org/package=rgbif>>.
- Colwell, R. K. and Rangel, T. F. 2009. Hutchinson's duality: the once and future niche. – *Proc. Natl Acad. Sci. USA* 106: 19651–19658.
- Dallas, T. et al. 2017. Species are not most abundant in the centre of their geographic range or climatic niche. – *Ecol. Lett.* 20: 1526–1533.
- Darroch, S. A. and Saupe, E. E. 2018. Reconstructing geographic range-size dynamics from fossil data. – *Paleobiology* 44: 25–39.
- De'Ath, G. 2007. Boosted trees for ecological modeling and prediction. – *Ecology* 88: 243–251.
- Elith, J. et al. 2008. A working guide to boosted regression trees. – *J. Anim. Ecol.* 77: 802–813.
- Fordyce, J. et al. 2016. Quantifying diet breadth through ordination of host association. – *Ecology* 97: 842–849.
- Friedman, J. H. 2002. Stochastic gradient boosting. – *Comput. Stat. Data Anal.* 38: 367–378.
- Gaston, K. J. and Blackburn, T. M. 1996. Range size-body size relationships: evidence of scale dependence. – *Oikos* 75: 479–485.
- Gaston, K. J. and Fuller, R. A. 2009. The sizes of species geographic ranges. – *J. Appl. Ecol.* 46: 1–9.
- Gaston, K. J. and Spicer, J. I. 2001. The relationship between range size and niche breadth: a test using five species of gammarus (amphipoda). – *Global Ecol. Biogeogr.* 10: 179–188.
- Graham, C. H. and Hijmans, R. J. 2006. A comparison of methods for mapping species ranges and species richness. – *Global Ecol. Biogeogr.* 15: 578–587.
- Greenwell, B. et al. 2022. *gbm*: generalized boosted regression models. R package ver. 2.1.8.1. – <<https://CRAN.R-project.org/package=gbm>>.
- Hijmans, R. 2021. *_geosphere: spherical trigonometry_*. R package ver. 1.5-14. – <<https://CRAN.R-project.org/package=geosphere>>.
- Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – *Int. J. Climatol.* 25: 1965–1978.
- Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. – *Am. Nat.* 163: 192–211.
- Jones, K. E. et al. 2009. Pantheria: a species-level database of life history, ecology and geography of extant and recently extinct mammals. – *Ecology* 90: 2648–2648.
- Kambach, S. et al. 2018. Of niches and distributions: range size increases with niche breadth both globally and regionally but regional estimates poorly relate to global estimates. – *Ecography* 42: 467–477.
- Kattge, J. et al. 2011. Try – a global database of plant traits. – *Global Change Biol.* 17: 2905–2935.
- Kriticos, D. J. et al. 2014. Extending the suite of bioclim variables: a proposed registry system and case study using principal components analysis. – *Methods Ecol. Evol.* 5: 956–960.
- Lichti, N. I. and Swihart, R. K. 2011. Estimating utilization distributions with kernel versus local convex hull methods. – *J. Wildl. Manage.* 75: 413–422.
- Machovsky-Capuska, G. E. et al. 2016. The multidimensional nutritional niche. – *Trends Ecol. Evol.* 31: 355–365.
- Mashau, A. C. et al. 2021. Plant height and lifespan predict range size in southern african grasses. – *J. Biogeogr.* 48: 3047–3059.
- Moore, T. E. et al. 2018. Spatial autocorrelation inflates niche breadth–range size relationships. – *Global Ecol. Biogeogr.* 27: 1426–1436.
- Morin, X. and Lechowicz, M. J. 2013. Niche breadth and range area in north american trees. – *Ecography* 36: 300–312.
- Morueta-Holme, N. et al. 2013. Habitat area and climate stability determine geographical variation in plant species range sizes. – *Ecol. Lett.* 16: 1446–1454.
- Myhrvold, N. P. et al. 2015. An amniote life-history database to perform comparative analyses with birds, mammals and reptiles. – *Ecology* 96: 3109–3109.
- Pagel, J. et al. 2020. Mismatches between demographic niches and geographic distributions are strongest in poorly dispersed and highly persistent plant species. – *Proc. Natl Acad. Sci. USA* 117: 3663–3669.
- Parnell, A. and Jackson, A. 2013. *siar*: stable isotope analysis in R. – <<https://cran.r-project.org/src/contrib/Archive/siar/>>.
- Pateiro-Lopez, B. et al. 2016. *alphahull*: generalization of the convex hull of a sample of points in the plane. R package ver. 2.5. – <<https://CRAN.R-project.org/package=alphahull>>.
- Pineda-Munoz, S. et al. 2021. Mammal species occupy different climates following the expansion of human impacts. – *Proc. Natl Acad. Sci.* 118: e1922859118.
- Pulliam, H. R. 2000. On the relationship between niche and distribution. – *Ecol. Lett.* 3: 349–361.
- Quinn, R. M. et al. 1996. Relative measures of geographic range size: empirical comparisons. – *Oecologia* 107: 179–188.
- Rolando, A. 1990. On niche breadth and related concepts. – *Ital. J. Zool.* 57: 145–148.
- Roslin, T. et al. 2017. Higher predation risk for insect prey at low latitudes and elevations. – *Science* 356: 742–744.
- Saupe, E. E. et al. 2019. Non-random latitudinal gradients in range size and niche breadth predicted by spatial patterns of climate. – *Global Ecol. Biogeogr.* 28: 928–942.
- Sheth, S. N. et al. 2020. Determinants of geographic range size in plants. – *New Phytol.* 226: 650–665.
- Slatyer, R. A. et al. 2013. Niche breadth predicts geographical range size: a general ecological pattern. – *Ecol. Lett.* 16: 1104–1114.
- Soberon, J. 2007. Grinnellian and eltonian niches and geographic distributions of species. – *Ecol. Lett.* 10: 1115–1123.
- Stevens, G. C. 1989. The latitudinal gradient in geographical range: how so many species coexist in the tropics. – *Am. Nat.* 133: 240–256.
- Sullivan, B. L. et al. 2009. *ebird*: a citizen-based bird observation network in the biological sciences. – *Biol. Conserv.* 142: 2282–2292.
- Swanson, H. K. et al. 2015. A new probabilistic method for quantifying n-dimensional ecological niches and niche overlap. – *Ecology* 96: 318–324.
- Vázquez, D. P. and Stevens, R. D. 2004. The latitudinal gradient in niche breadth: concepts and evidence. – *Am. Nat.* 164: E1–E19.
- Violle, C. and Jiang, L. 2009. Towards a trait-based quantification of species niche. – *J. Plant Ecol.* 2: 87–93.
- Woudenberg, S. W. et al. 2010. The forest inventory and analysis database: database description and users manual ver. 4.0 for phase 2. Gen. Tech. Rep. RMRS-GTR-245. – US Dept of Agriculture, Forest Service, Rocky Mountain Research Station.
- Yu, F. et al. 2017. Climatic niche breadth can explain variation in geographical range size of alpine and subalpine plants. – *Int. J. Geogr. Inform. Sci.* 31: 190–212.