

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

Tad A Dallas^{a,b,*} and Andrew Kramer^c

^a*Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA*

^b*Department of Biological Sciences, University of South Carolina, Columbia, SC, USA*

^c*Department of Integrative Biology, University of South Florida, Tampa, FL, USA*

*Corresponding author: tad.a.dallas@gmail.com

Author contributions: TAD conceptualized the study and performed the analysis. TAD wrote the first draft and AMK contributed to later revisions of the analysis and text. All authors approved submission of the manuscript.

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. 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. Discussions with Laura Antao greatly improved the manuscript. The study has been supported by the TRY initiative on plant traits (<http://www.trynodb.org>). Silhouette images used in figures are from Phylopic (<http://phylopic.org/>).

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

Keywords: Rapoport's rule, Convex hull, Species range estimation, Geographic range size, Latitudinal gradient

Words in text: 4750

Words outside text (captions, etc.): 500

Number of tables: 1 (main text), 0 (supplement)

Number of figures: 5 (main text), 16 (supplement)

Number of cited references: 53

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

3 **Abstract**

4 Species with broader niches may have the opportunity to occupy larger geographic
5 areas, assuming no limitations on dispersal and a relatively homogeneous envi-
6 ronmental space. Here, we use data on a large set of mammal ($n = 1225$), bird
7 ($n = 1829$), and tree ($n = 341$) species to examine the **1**) relationship between
8 geographic range size and climatic niche area, **2**) influence of species traits on
9 species departures from this relationship, and **3**) sensitivity of these relationships
10 to how species range size and climatic niche area are estimated. We find positive
11 *geographic range size – climatic niche area* relationships for all taxa, with residual
12 variation depending on latitude but not differing from a null model. Together, we
13 provide support for this general macroecological relationship which is dependent
14 on space, but not on species traits, and no different from a null model.

15 **Running title:** Geographic range and climatic niche

16

17 Introduction

18 Geographically widespread species tend to also have larger ecological niches rela-
19 tive to smaller-ranged species (Brown, 1984, Slatyer *et al.*, 2013, Yu *et al.*, 2017).
20 Niche space can be defined in terms of niche breadth (i.e., the range of resources
21 that a species uses Rolando (1990)) or climatic niche area (i.e., the range of cli-
22 matic conditions the species occurs in Dallas *et al.* (2017)). Previous investigations
23 into the relationship between species geographic range size and niche size have de-
24 fined the niche in several ways, including the use of habitat (e.g., number biomes
25 occupied), diet (e.g., number of different food types used), and environmental tol-
26 erance (e.g., elevational range) measures (Slatyer *et al.*, 2013). The difficulties in
27 defining and quantifying species geographic range (Gaston & Fuller, 2009) and
28 climatic niche (Fordyce *et al.*, 2016, Machovsky-Capuska *et al.*, 2016, Rolando,
29 1990, Violle & Jiang, 2009) size have contributed to the mixed support for these
30 scaling relationships (Gaston & Spicer, 2001, Morueta-Holme *et al.*, 2013, Slatyer
31 *et al.*, 2013). However, two fairly recent studies found strong support across many
32 different species for positive geographic range size - niche size relationships (Kam-
33 bach *et al.*, 2018, Slatyer *et al.*, 2013), suggesting that these relationships may be
34 quite general, but sensitive to spatial scale (Kambach *et al.*, 2018).

35 Examining the relationship between geographic range size and climatic niche
36 area also affords another interesting avenue; examining the effects of different
37 geographic range size and niche area estimation procedures on subsequent scaling

relationships. With respect to geographic range size estimation, Gaston & Fuller (2009) differentiated area of occupancy (A00) from extent of occurrence (E00), where A00 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 E00. A consensus on best practices for estimation of species geographic range size (Gaston & Fuller, 2009, Graham & Hijmans, 2006) 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 & 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.

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. 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

59 climatic niches will be capable of colonizing a larger set of geographic locations
60 (Pulliam, 2000). Disentangling a true relationship between an artifact of a spatial
61 sampling process is difficult with observational data, as the niche and the geo-
62 graphic distribution are inherently linked through the observation process (Colwell
63 & Rangel, 2009, Pulliam, 2000).

64 From a practical perspective, the potential uncoupling of geographic range size
65 and climatic niche area as a function of spatial processes can lead to species di-
66 verging from the expected global relationship, which would consider data on all
67 available species within some taxonomic group(s) to estimate the relationship be-
68 tween geographic range size and climatic niche area. However, the availability of
69 colonizable land and climatic niche area is not uniformly distributed across space,
70 suggesting that spatial gradients can constrain geographic range size – climatic
71 niche area relationships even when all species are functionally neutral (Blackburn
72 & Gaston, 1997). This creates an interesting possibility; the availability of land
73 area and climatic niche space in a given area places fundamental constraints on
74 the resulting geographic range size – climatic niche area relationship. This could
75 suggest the existence of a spatial signal in divergence from the global geographic
76 range size – climatic niche area relationship, potentially driven by geographic areas
77 of high discordance (e.g., large geographic space with low climatic heterogeneity).

78 But while a geographic range size – climatic niche area relationship might vary
79 across spatial gradients and be expected under neutral assumptions, divergence
80 from the expected scaling relationship across a wide number of species might point
81 to an underlying species attribute associated with divergence (e.g., dispersal ability
82 (Pagel *et al.*, 2020)). Relating the residual deviations from the expected relation-
83 ship to geographic covariates and species traits may provide insight into when and
84 where relationships between species geographic range size and climatic niche area
85 are weakened. For instance, such a process could identify species with large geo-
86 graphic ranges with smaller than expected niche areas. This would tend to occur
87 in geographic locations which remain climatically similar across large geographic
88 spaces. That is, a species may specialize on a very common set of environments,
89 leading to a rather small climatic niche area and a large geographic range size.

90 Here, we use data on a large set of mammal ($n = 1225$), bird ($n = 1829$), and
91 tree ($n = 341$) species distributed across the Americas to examine the **1**) relation-
92 ship between geographic range size and climatic niche area, **2**) influence of species
93 traits on species departures from the best fit geographic range size – climatic niche
94 area relationship, and **3**) sensitivity to different species geographic range size or
95 climatic niche area estimation approaches. The relationship between geographic
96 range size and climatic niche area appears to be quite general, likely as a result
97 of the close relationship between geographic distribution and the niche (Pulliam,
98 2000). The latitudinal centre of a species geographic range was the best predictor

99 of the residual variation from the geographic range size – climatic niche area re-
100 lationship, though species body mass was also an important predictor, providing
101 evidence for a clear spatial signal in geographic range size – climatic niche area
102 relationships. We highlight the latitudinal variation in potential colonizable land
103 area and climatic niche space, demonstrating that this fundamentally constrains
104 potential relationships. Residual variation in geographic range size – climatic niche
105 area relationships did not strongly differ from a null model, suggesting aspects of
106 the landscape may drive species deviations from a global relationship to a much
107 greater extent than species trait variation. Together, we provide strong evidence
108 for the generality of geographic range size – climatic niche area relationships, high-
109 lighting the role of potential fundamental constraints on the relationship given the
110 spatial distribution of land area and climatic niche space.

111 **Methods**

112 **Data sources**

113 We obtained species occurrence and trait data for mammal, tree, and bird species
114 from freely available data sources (described below). To include as many species
115 as possible, we queried species occurrence records from the Global Biodiversity
116 Information Facility – a species occurrence database – using the R package `rgbif`
117 (Chamberlain *et al.*, 2016) for all mammal species listed in PanTHERIA (Jones

118 *et al.*, 2009), a mammal trait database. Species occurrence records were checked
 119 for quality using **scrubr**, which removed occurrences with missing, uncertain, or
 120 unlikely (0 °N 0 °W) latitude and longitude values (Chamberlain, 2016). We
 121 obtained bird occurrence data from eBird (Sullivan *et al.*, 2009), and bird trait
 122 data from Myhrvold *et al.* (2015). Lastly, tree occurrence data was obtained
 123 from the US Department of Agriculture Forest Inventory and Analysis database
 124 (Bechtold *et al.*, 2005), and tree trait data was obtained from TRY (Kattge *et al.*,
 125 2011). Relevant species traits are described below (and also in Table 1), but
 126 largely center on life history traits influencing the range of habitats a species can
 127 occupy or the pace of life (e.g., birth rate) of a given species, taken as a subset from
 128 available trait data sources. Trait data for some species were unavailable, resulting
 129 in variable trait data coverage (e.g., 90% coverage on mammal body sizes, but only
 130 around 30% coverage on mammal lifespan).

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

139 included in the analyses.

140 **Estimation of geographic range and climatic niche area**

141 Many methods have been developed to estimate species geographic range and
142 climatic niche area (Burgman & Fox, 2003, Lichi & Swihart, 2011, Quinn *et al.*,
143 1996), each of which makes tacit assumptions about the structure of the climatic
144 niche or the spatial distribution of a species across a landscape (Figure 1). For
145 instance, the convex hull approach may potentially overestimate geographic range
146 area as a result of the limitation that the minimum bounding polygon of the species
147 range can only contain convex angles, leading to geographic areas which may
148 contain regions of geographic space where a species cannot exist. More restrictive
149 approaches – such as alpha hulls – attempt to account for this, but, in doing so,
150 may (Darroch & Saupe, 2018) or may not (Burgman & Fox, 2003) be more prone
151 to issues with sampling or detection bias. There is presently no clear optimal
152 method for estimating species geographic range sizes or climatic niche areas. As
153 such, we use three well-established methods; convex hull, alpha hull (with two
154 different alpha parameterizations) using the *alphahull* R package (Pateiro-Lopez
155 *et al.*, 2016), and standard ellipse area using the *siar* R package (Parnell & Jackson,
156 2013). In the main text, we estimate species ranges using the convex hull, as it
157 is well-established and does not require parameterization like other methods (e.g.,
158 α parameter of alpha hulls). In the Supplement, we discuss the sensitivity of the

159 scaling relationship between geographic range size and climatic niche area to the
160 measure used to estimate geographic and climatic niche areas.

161 We operationalize the species niche as the set of climatic space a species occupies
162 (most akin to a realized niche (Soberon, 2007)). Climatic niche area was deter-
163 mined by first translating the multivariate climate space into a two-dimensional
164 space comparable to geographic space. To do this, we calculated the first two
165 principal components (PCA) of the set of 56 BioClim/WorldClim variables (Hij-
166 mans *et al.*, 2005), projecting geographic coordinates to climatic niche space and
167 explaining over 77% of the total global climatic variation in a two-dimensional
168 space (Kambach *et al.*, 2018, Kriticos *et al.*, 2014). The WorldClim variables (n
169 = 36), containing monthly information on minimum and maximum temperature
170 and precipitation, and the BioClim variables ($n = 19$), containing derived quanti-
171 ties such as temperature seasonality and mean annual precipitation, represent the
172 best available large scale climatic data currently (Barbet-Massin & Jetz, 2014). In
173 terms of describing the species niche, this approach reduces the global climate to
174 a small number of dimensions, and may not capture species-specific niche require-
175 ments. For instance, hibernating species may response more strongly to climatic
176 conditions in spring instead of in winter. However, by compressing monthly tem-
177 perature and precipitation data down to two axes we have explained the majority
178 of relevant climatic variation – at least for large-scale patterns across species entire
179 geographic ranges. By examining the scree plot of the first two PCA axes of the

180 global climatic space, it is clear that the first PCA axis largely corresponds to tem-
181 perature covariates, while the second PCA axis corresponds more to precipitation
182 covariates (Figure S2).

183 **Latitudinal variation in potential geographic range size –** 184 **climatic niche area relationship**

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

193 **Geographic and species traits associated with residuals**

194 We related square-root-transformed species geographic range size to square-root-
195 transformed species climatic niche area using linear regression (Figure S3). The
196 residuals from these simple linear models represent the divergence of each species
197 from the overall relationship between geographic range size and climatic niche
198 area. If this relationship produces residuals that are no different from a null model,

199 this would suggest that the residual variation is more a function of spatial auto-
200 correlation, sampling and measurement error, or differential species distributions
201 independent of species identity.

202 To examine the null distribution of residuals, we simulated species geographic
203 distributions and climatic niches by sampling the empirical occurrence values
204 across ranges of latitude, total latitudinal range, and occurrence number that were
205 comparable to the empirical data (see Figure S4). That is, we selected a random
206 latitudinal minimum ($Uniform(-57, 57)$ for mammals and birds, $Uniform(18, 48)$
207 for trees to match their empirical distributions), a random latitudinal range size
208 ($Uniform(1, 45)$), and a random number of occurrence points ($Uniform(5, 500)$).
209 Given these ranges, we assembled species distributions by sampling the empirical
210 occurrence data for mammals, trees, and birds separately. This was performed
211 5000 times for each species group, then subsampled down to the number of species
212 in each taxa to allow for a more direct comparison. By comparing the distribution
213 of residual values from both empirical and null geographic range size – climatic
214 niche area relationships, we explore to what extent this macroecological pattern is
215 simply an emergent property of spatially autocorrelated environmental data, or if
216 geographic or trait variation can drive departures from the expected relationship.

217 We then related these residuals to geographic variables and species traits using
218 an ensemble regression tree approach. To examine spatial structure of residual

219 variation in geographic range size – climatic niche area relationships, we included
220 the latitudinal centre of a species range as a covariate. With respect to species
221 traits, we selected life history traits (e.g., body size, habitat breadth, trophic
222 level, dispersal distance) that have been previously found to be related to species
223 geographic range size (Table 1). Further, we also selected some traits related to
224 species ‘pace of life’ or demography, including litter size, lifespan, and seed mass
225 (Table 1).

226 Gradient boosted machines, also known as boosted regression trees, are a flexi-
227 ble regression technique in which many weak learning decision trees are iteratively
228 created, where each tree attempts to explain variation left over from the previous
229 tree (Elith *et al.*, 2008, Friedman, 2002). When these trees are combined, they are
230 able to handle collinear data, handle missing values, and account for variable in-
231 teractions (De’Ath, 2007, Elith *et al.*, 2008, Friedman, 2002). Models were trained
232 in *R* using the *gbm* package (Ridgeway, 2017), and were internally five-fold cross
233 validated on 80% of the data to avoid overfitting, while the remaining 20% test
234 data was used to evaluate model performance. We trained 10 models per species
235 group, each on a different random subset of 80% of the data, in order to exam-
236 ine variation in model performance as a function of the sampled data. Model
237 performance was assessed using Spearman’s rank correlation coefficients between
238 residuals from the relationship between geographic range size and climatic niche
239 area that were predicted from the boosted regression tree model compared to the

240 actual values observed in the test data.

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

249 Data and *R* code to reproduce the results is provided at
250 <https://doi.org/10.6084/m9.figshare.7964666.v2>.

251 Results

252 The relationship between geographic range and climatic niche 253 area

254 Significantly positive relationships were observed between geographic range area
255 and climatic niche area (Figure 2) for mammals ($\beta = 4.21$, $p < 0.0001$, adjusted
256 $R^2 = 0.44$), trees ($\beta = 2.77$, $p < 0.0001$, adjusted $R^2 = 0.41$), and birds ($\beta =$
257 4.72 , $p < 0.0001$, adjusted $R^2 = 0.60$). This supports previous findings suggest-

ing the generality of this relationship (Kambach *et al.*, 2018, Slatyer *et al.*, 2013). Residuals from this linear relationship were not detectably different in their distribution from a null expectation generated from sampling empirical occurrence data for randomly selected latitudinal bands and number of occurrence points (Figure 3). This does not mean that geographic, community, and trait variables are unimportant to geographic range size – climatic niche area relationships, but simply that the distribution of the residuals from the empirical relationship, which was markedly positive, do not appear different from what would be expected from a simple null model. This 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 (Figure S14) or climatic niche area (Figure S15) resulted in more weakly related estimates (see Supplemental Materials; Figures S5 - S13). Across a latitudinal gradient in the Americas, the potential relationship space between geographic range size and climatic niche area varied drastically (Figure 4), suggesting that the slope of the relationship may be determined – at least in part – by the latitudinal range where the species is found.

276 **Geographic and species traits associated with residuals**

277 Residuals from the relationship between species geographic range size and climatic
278 niche area were fairly well-predicted by species traits (Table 1) for mammal ($\bar{\rho} =$
279 0.53 ± 0.07), tree ($\bar{\rho} = 0.80 \pm 0.02$), and bird ($\bar{\rho} = 0.55 \pm 0.03$) species (Fig-
280 ure 5). Latitudinal centre was the most important covariate in boosted regression
281 tree models for all species groups, suggesting a strong latitudinal signal on the
282 geographic range size – climatic niche area relationship. For tree and bird species,
283 the importance of latitudinal range centre dominated the model performance, with
284 the remaining species traits contributing little to model performance (Figure 5).
285 However, for mammal species, species body mass was important for model perfor-
286 mance (Figure 5), potentially because of the positive relationship between latitude
287 and body size (i.e., Bergmann’s rule; Ashton *et al.* (2000)). Evidence for this
288 comes from the partial dependence plots, which examine the relative effect of each
289 covariate on the residual variation in the geographic range size – climatic niche
290 area relationship (Figure S1).

291 The importance of latitudinal centre to residual variation in geographic range
292 size – climatic niche area relationships is not a function of latitudinal structure in
293 either geographic range or climatic niche across species (see Supplemental Figure
294 S16), suggesting that the relationship between climatic niche area and geographic
295 range size is influenced by latitude, but both area estimates are not strongly related
296 to species latitudinal centre independently. Dividing the Americas into latitudi-

297 nal bands, it becomes apparent that there is a latitudinal signal in the potential
298 geographic range size – climatic niche area relationships that could emerge for a
299 given species with a contiguous range (Figure 4). This does not inherently sug-
300 gest that the residual variation in the geographic range size – climatic niche area
301 relationship will be latitudinally structured, but simply that the distribution of
302 potential climatic niche space and available land area does contain a latitudinal
303 signal (Figures 1 and 4), which may underlie our failure to detect a difference in
304 the residual variation in the empirical data and the null model.

305 A positive effect of latitudinal range centre suggests that species distributed
306 at higher absolute latitudes tend to have larger climatic niches than expected
307 given their geographic range sizes. However, given that the modeling approach
308 used accounts for variable interactions and non-linear relationships, we see clear
309 non-linear relationships between latitudinal range centre and residuals from the
310 geographic range size – climatic niche area relationship, with clear differences
311 among taxa as well (Figure S1). A negative relationship between mammal body
312 mass and the residual variation in the geographic range size – climatic niche area
313 relationship suggests that larger-bodied mammals tend to have smaller climatic
314 niches than expected given their geographic range size (Figure S1). We see clear
315 differences in the effect of body mass on residual variation of the geographic range
316 size – climatic niche area relationship between mammals (negative relationship)
317 and birds (positive relationship) (Figure S1). These relationships were further

318 supported when estimating geographic range size and climatic niche area using
319 standard ellipse areas (SIAR), though the use of alpha hulls resulted in reduced
320 model performance (see Supplemental Materials).

321 Discussion

322 Geographic range size was positively related to climatic niche area for all species
323 groups, supporting previous studies claiming this is a consistent ecological re-
324 lationship (reviewed in Slatyer *et al.* (2013)). The scaling between geographic
325 range size and climatic niche area may stem from the close relationship between a
326 species niche and its corresponding distribution (Pulliam, 2000), though this tac-
327 itly assumes that species with broader climatic niches will inherently have larger
328 geographic range areas, owing in part to the commonly observed spatial auto-
329 correlation in climatic conditions. The assumed relationship between geographic
330 range size and climatic niche area was supported by both our analyses and the
331 null model, which generated residual variation which matched the empirical data.
332 However, we were also able to detect a latitudinal signal in the residual variation
333 of the geographic range size – climatic niche area relationship. While this spatial
334 structure appeared to take different forms for the 3 groups of species, the impor-
335 tance of latitude was likely a function of the distribution of available land area
336 and climatic niche space, caused by the latitudinal decoupling between available
337 land area and climatic niche space. Together, our findings suggest that geographic

338 range size positively scales with climatic niche area, and that the residual variation
339 in this relationship may be explained by species life history traits and geographic
340 position.

341 The existence of latitudinal structure in species deviations from a general geo-
342 graphic range size – climatic niche area relationship is an important finding, indi-
343 cating a clear geographic constraint that outweighs a suite of species traits. The
344 existence of spatial structure in available geographic range size and climatic niche
345 area is an important consideration as well (Figure 4), as available land area and
346 niche area could fundamentally constrain the geographic range size – climatic niche
347 relationship along spatial gradients. Further, there is mixed evidence suggesting
348 that geographic range size (Stevens, 1989) and climatic niche breadth (Vázquez
349 & Stevens, 2004) should scale positively with latitude. We did not find strong
350 evidence for latitudinal scaling in square-root-transformed geographic range size
351 or climatic niche area, where latitude was defined using the latitudinal centre for
352 each species (see Supplemental Figure S16). However, we did see a clear increase
353 in available geographic space in northern latitudes, while available niche space is
354 highest in the tropics (Figure 1). This creates a situation where the available ge-
355 ographic area and climatic niche space are latitudinally structured, constraining
356 the possible relationship between geographic range size and climatic niche area for
357 a given species occupying a contiguous geographic range (Figure 4). Other meth-
358 ods of range and niche estimation may partially remove this latitudinal structure

359 (e.g., alpha hulls which allow discontinuous ranges and niches), though latitudinal
360 centre tended to still be important when using alpha hulls here (see Supplemental
361 Materials).

362 Apart from the spatial distribution of available land area and climatic niche
363 space, many other relevant factors follow a latitudinal gradient. For instance,
364 species diversity (Hillebrand, 2004) and ecological interactions (Roslin *et al.*, 2017)
365 can be latitudinally-structured, and the latitudinally-structured residual variation
366 in geographic range size–climatic niche area relationships may be a function of
367 competition, natural enemies, or dispersal limitation (Pagel *et al.*, 2020). Further,
368 geological and evolutionary processes underlying the distribution of species diver-
369 sity and the shape of continents may further influence the resulting variation in
370 species geographic range size – climatic niche area relationships. Understanding
371 the relative roles of species interactions, species traits, and the entirely neutral
372 constraints of available land area and climatic niche space is an important next
373 step.

374 We recognize that increasing geographic range size is likely to increase climatic
375 niche area, as niche axes often contain a strong spatial signal. This could lead to
376 reduced niche breadth in the tropics solely as a result of the spatial structure of
377 the available climatic niche space (Saupe *et al.*, 2019). However, by focusing on ex-
378 plaining 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. 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.

References

- Addo-Bediako, A., Chown, S.L. & Gaston, K.J. (2000). Thermal tolerance, climatic variability and latitude. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267, 739–745.
- Ashton, K.G., Tracy, M.C. & Queiroz, A.d. (2000). Is bergmanns rule valid for mammals? *The American Naturalist*, 156, 390–415.
- Barbet-Massin, M. & Jetz, W. (2014). A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Diversity and Distributions*, 20, 1285–1295.
- Bechtold, W.A., Patterson, P.L. *et al.* (2005). The enhanced forest inventory and analysis program-national sampling design and estimation procedures. *Gen. Tech. Rep. SRS-80. Asheville, NC: US Department of Agriculture, Forest Service, Southern Research Station. 85 p.*, 80.
- Blackburn, T.M. & Gaston, K.J. (1997). The relationship between geographic area and the latitudinal gradient in species richness in New World birds. *Evolutionary Ecology*, 11, 195–204.
- Blonder, B., Lamanna, C., Violle, C. & Enquist, B.J. (2014). The n-dimensional hypervolume. *Global Ecology and Biogeography*, 23, 595–609.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5–32.

- 441 Brown, J.H. (1984). On the relationship between abundance and distribution of
 442 species. *The american naturalist*, 124, 255–279.
- 443 Burgman, M.A. & Fox, J.C. (2003). Bias in species range estimates from mini-
 444 mum convex polygons: implications for conservation and options for improved
 445 planning. In: *Animal Conservation forum*. Cambridge University Press, vol. 6,
 446 pp. 19–28.
- 447 Chamberlain, S. (2016). *scrubr: Clean Biological Occurrence Records*. R package
 448 version 0.1.1.
- 449 Chamberlain, S., Ram, K., Barve, V. & McGlinn, D. (2016). *rgbif: Interface to the*
 450 *Global 'Biodiversity' Information Facility 'API'*. R package version 0.9.2.9430.
- 451 Colwell, R.K. & Rangel, T.F. (2009). Hutchinson’s duality: the once and future
 452 niche. *Proceedings of the National Academy of Sciences*, 106, 19651–19658.
- 453 Dallas, T., Decker, R.R. & Hastings, A. (2017). Species are not most abundant
 454 in the centre of their geographic range or climatic niche. *Ecology letters*, 20,
 455 1526–1533.
- 456 Darroch, S.A. & Saupe, E.E. (2018). Reconstructing geographic range-size dy-
 457 namics from fossil data. *Paleobiology*, 44, 25–39.
- 458 De’Ath, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology*,
 459 88, 243–251.

- 460 Elith, J., Leathwick, J.R. & Hastie, T. (2008). A working guide to boosted regres-
461 sion trees. *Journal of Animal Ecology*, 77, 802–813.
- 462 Fordyce, J., Nice, C., Hamm, C. & Forister, M.L. (2016). Quantifying diet breadth
463 through ordination of host association. *Ecology*.
- 464 Friedman, J.H. (2002). Stochastic gradient boosting. *Computational Statistics &*
465 *Data Analysis*, 38, 367–378.
- 466 Gaston, K.J., Blackburn, T.M. & Spicer, J.I. (1998). Rapoport’s rule: time for an
467 epitaph? *Trends in Ecology & Evolution*, 13, 70–74.
- 468 Gaston, K.J. & Fuller, R.A. (2009). The sizes of species geographic ranges. *Journal*
469 *of applied ecology*, 46, 1–9.
- 470 Gaston, K.J. & Spicer, J.I. (2001). The relationship between range size and niche
471 breadth: a test using five species of gammarus (amphipoda). *Global Ecology*
472 *and Biogeography*, 10, 179–188.
- 473 Graham, C.H. & Hijmans, R.J. (2006). A comparison of methods for mapping
474 species ranges and species richness. *Global Ecology and biogeography*, 15, 578–
475 587.
- 476 Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005). Very
477 high resolution interpolated climate surfaces for global land areas. *International*
478 *Journal of Climatology*, 25, 1965–1978.

- 479 Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *The*
480 *American Naturalist*, 163, 192–211.
- 481 Jones, K.E., Bielby, J., Cardillo, M., Fritz, S.A., O’Dell, J., Orme, C.D.L., Safi, K.,
482 Sechrest, W., Boakes, E.H., Carbone, C. *et al.* (2009). Pantheria: a species-level
483 database of life history, ecology, and geography of extant and recently extinct
484 mammals. *Ecology*, 90, 2648–2648.
- 485 Kambach, S., Lenoir, J., Decocq, G., Welk, E., Seidler, G., Dullinger, S., Gégout,
486 J.C., Guisan, A., Pauli, H., Svenning, J.C. *et al.* (2018). Of niches and distribu-
487 tions: range size increases with niche breadth both globally and regionally but
488 regional estimates poorly relate to global estimates. *Ecography*.
- 489 Kattge, J., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönsch, G., Garnier,
490 E., Westoby, M., Reich, P.B., Wright, I.J. *et al.* (2011). Try—a global database
491 of plant traits. *Global Change Biology*, 17, 2905–2935.
- 492 Kriticos, D.J., Jarošik, V. & Ota, N. (2014). Extending the suite of bioclim
493 variables: a proposed registry system and case study using principal components
494 analysis. *Methods in Ecology and Evolution*, 5, 956–960.
- 495 Lichti, N.I. & Swihart, R.K. (2011). Estimating utilization distributions with
496 kernel versus local convex hull methods. *The Journal of Wildlife Management*,
497 75, 413–422.
- 498 Machovsky-Capuska, G.E., Senior, A.M., Simpson, S.J. & Raubenheimer, D.

499 (2016). The multidimensional nutritional niche. *Trends in ecology & evolution*,
 500 31, 355–365.

501 Morueta-Holme, N., Enquist, B.J., McGill, B.J., Boyle, B., Jørgensen, P.M., Ott,
 502 J.E., Peet, R.K., Šímová, I., Sloat, L.L., Thiers, B. *et al.* (2013). Habitat area
 503 and climate stability determine geographical variation in plant species range
 504 sizes. *Ecology letters*, 16, 1446–1454.

505 Myhrvold, N.P., Baldrige, E., Chan, B., Sivam, D., Freeman, D.L. & Ernest,
 506 S.M. (2015). An amniote life-history database to perform comparative analyses
 507 with birds, mammals, and reptiles. *Ecology*, 96, 3109–3109.

508 Pagel, J., Treurnicht, M., Bond, W.J., Kraaij, T., Nottebrock, H., Schutte-Vlok,
 509 A., Tonnabel, J., Esler, K.J. & Schurr, F.M. (2020). Mismatches between de-
 510 mographic niches and geographic distributions are strongest in poorly dispersed
 511 and highly persistent plant species. *Proceedings of the National Academy of*
 512 *Sciences*, 117, 3663–3669.

513 Parnell, A. & Jackson, A. (2013). *siar: Stable Isotope Analysis in R*. R package
 514 version 4.2.

515 Pateiro-Lopez, B., Rodriguez-Casal, A. & . (2016). *alphahull: Generalization of*
 516 *the Convex Hull of a Sample of Points in the Plane*. R package version 2.1.

517 Pulliam, H.R. (2000). On the relationship between niche and distribution. *Ecology*
 518 *letters*, 3, 349–361.

519 Quinn, R.M., Gaston, K.J. & Arnold, H.R. (1996). Relative measures of geographic
520 range size: empirical comparisons. *Oecologia*, 107, 179–188.

521 Ridgeway, G. (2017). *gbm: Generalized Boosted Regression Models*. R package
522 version 2.1.3.

523 Rohde, K. (1996). Rapoport’s rule is a local phenomenon and cannot explain
524 latitudinal gradients in species diversity. *Biodiversity Letters*, pp. 10–13.

525 Rolando, A. (1990). On niche breadth and related concepts. *Italian Journal of*
526 *Zoology*, 57, 145–148.

527 Roslin, T., Hardwick, B., Novotny, V., Petry, W.K., Andrew, N.R., Asmus, A.,
528 Barrio, I.C., Basset, Y., Boesing, A.L., Bonebrake, T.C. *et al.* (2017). Higher
529 predation risk for insect prey at low latitudes and elevations. *Science*, 356,
530 742–744.

531 RoY, K., Jablonski, D. & Valentine, J.W. (1994). Eastern pacific molluscan
532 provinces and latitudinal diversity gradient: no evidence for" rapoport’s rule".
533 *Proceedings of the National Academy of Sciences*, 91, 8871–8874.

534 Saupe, E.E., Myers, C.E., Peterson, A.T., Soberón, J., Singarayer, J., Valdes, P.
535 & Qiao, H. (2019). Non-random latitudinal gradients in range size and niche
536 breadth predicted by spatial patterns of climate. *Global Ecology and Biogeogra-*
537 *phy*.

- 538 Slatyer, R.A., Hirst, M. & Sexton, J.P. (2013). Niche breadth predicts geographical
539 range size: a general ecological pattern. *Ecology letters*, 16, 1104–1114.
- 540 Soberon, J. (2007). Grinnellian and eltonian niches and geographic distributions
541 of species. *Ecology Letters*, 10, 1115–1123.
- 542 Stevens, G.C. (1989). The latitudinal gradient in geographical range: how so many
543 species coexist in the tropics. *The American Naturalist*, 133, 240–256.
- 544 Stevens, G.C. (1992). The elevational gradient in altitudinal range: an extension of
545 rapoport’s latitudinal rule to altitude. *The American Naturalist*, 140, 893–911.
- 546 Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D. & Kelling, S. (2009).
547 ebird: A citizen-based bird observation network in the biological sciences. *Bio-*
548 *logical Conservation*, 142, 2282–2292.
- 549 Swanson, H.K., Lysy, M., Power, M., Stasko, A.D., Johnson, J.D. & Reist, J.D.
550 (2015). A new probabilistic method for quantifying n-dimensional ecological
551 niches and niche overlap. *Ecology*, 96, 318–324.
- 552 Vázquez, D.P. & Stevens, R.D. (2004). The latitudinal gradient in niche breadth:
553 concepts and evidence. *The American Naturalist*, 164, E1–E19.
- 554 Violle, C. & Jiang, L. (2009). Towards a trait-based quantification of species niche.
555 *Journal of Plant Ecology*, 2, 87–93.

- 556 Woudenberg, S.W., Conkling, B.L., OConnell, B.M., LaPoint, E.B., Turner, J.A.
557 & Waddell, K.L. (2010). The forest inventory and analysis database: Database
558 description and users manual version 4.0 for phase 2. *Gen. Tech. Rep. RMRS-*
559 *GTR-245. Fort Collins, CO: US Department of Agriculture, Forest Service,*
560 *Rocky Mountain Research Station.*
- 561 Yu, F., Groen, T.A., Wang, T., Skidmore, A.K., Huang, J. & Ma, K. (2017).
562 Climatic niche breadth can explain variation in geographical range size of alpine
563 and subalpine plants. *International journal of geographical information science,*
564 31, 190–212.

565 **Data accessibility**

566 Data and *R* code to reproduce the analyses is available on figshare at
567 <https://doi.org/10.6084/m9.figshare.7964666.v2>. Data are available for
568 eBird data (Sullivan *et al.*, 2009), Forest Inventory and Analysis data (Woudenberg
569 *et al.*, 2010) (<https://www.fia.fs.fed.us/>), and the Global Biodiversity Information
570 Facility data (Chamberlain *et al.*, 2016, Jones *et al.*, 2009). Authors should cite
571 the original data sources.

572

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	m
	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

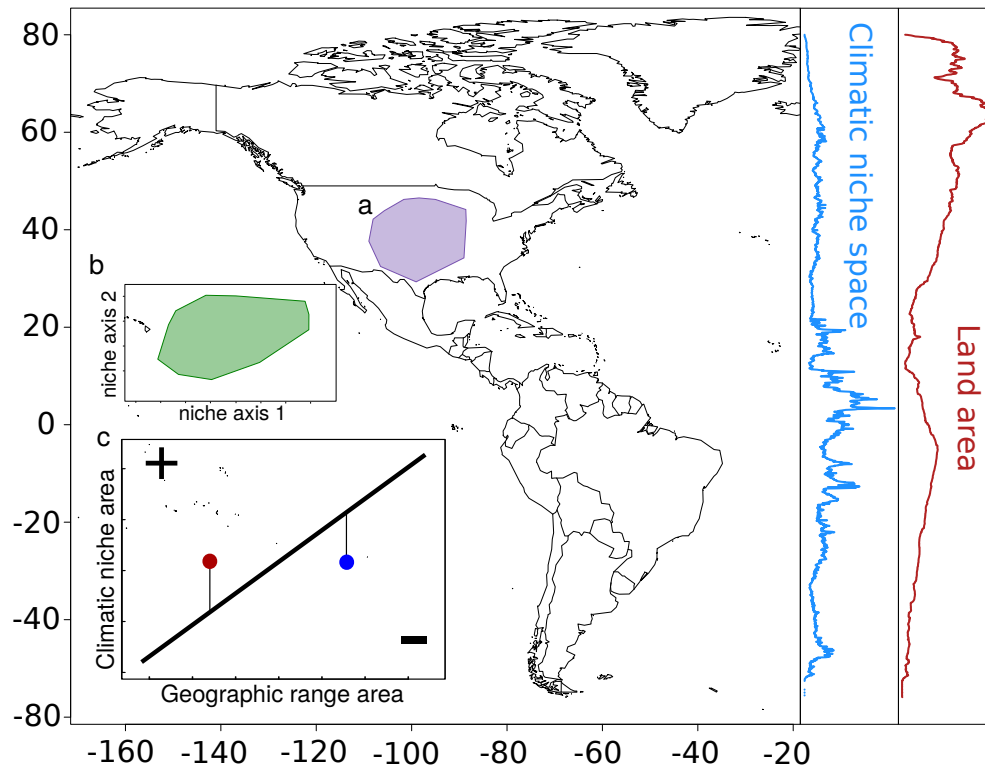


Figure 1: Residual variation from the relationship between a species geographic range size (spatial polygon in panel a) and corresponding climatic niche area (polygon in panel b) may be associated with species traits or spatial structure. These residuals (depicted in panel c) represent situations where niche area is larger than expected given the geographic range size (indicated with a red dot and a "+" symbol) or where geographic area is larger 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).

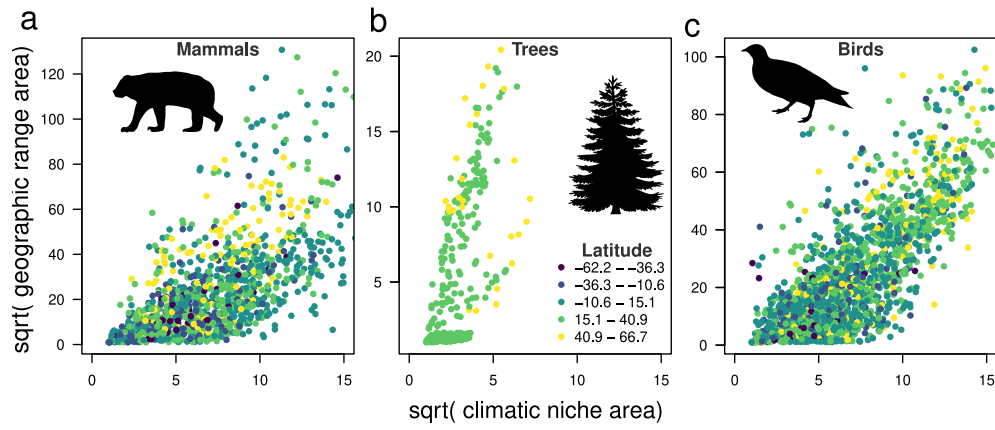


Figure 2: The relationship between a species geographic range size (square-root-transformed) and climatic niche area (square-root-transformed) for a set of 3395 species of a) mammals, b) trees, and c) birds. Point color is based on species' latitudinal centroids.

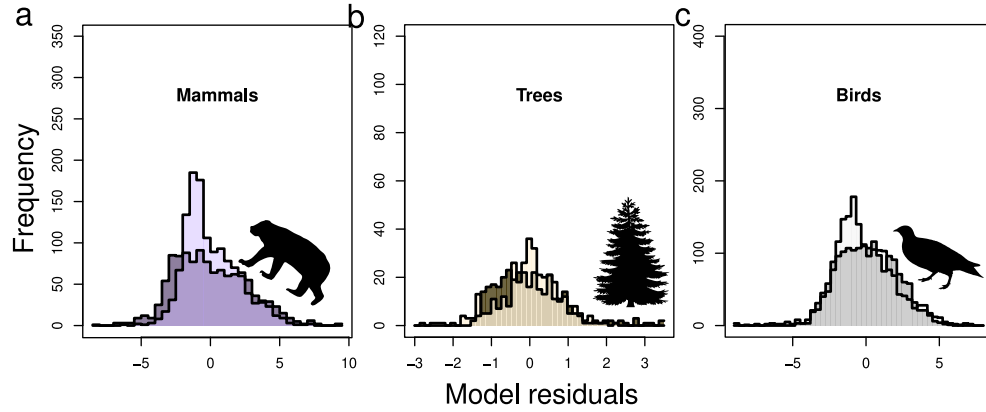


Figure 3: Empirical (darker colors) and null (lighter colors) distributions of residuals for a) mammal, b) tree, and c) bird species, suggesting that the residual variation observed in natural systems is quite similar to that observed from an uninformed null model which ignores interspecific interactions and other ecologically relevant processes. While 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.

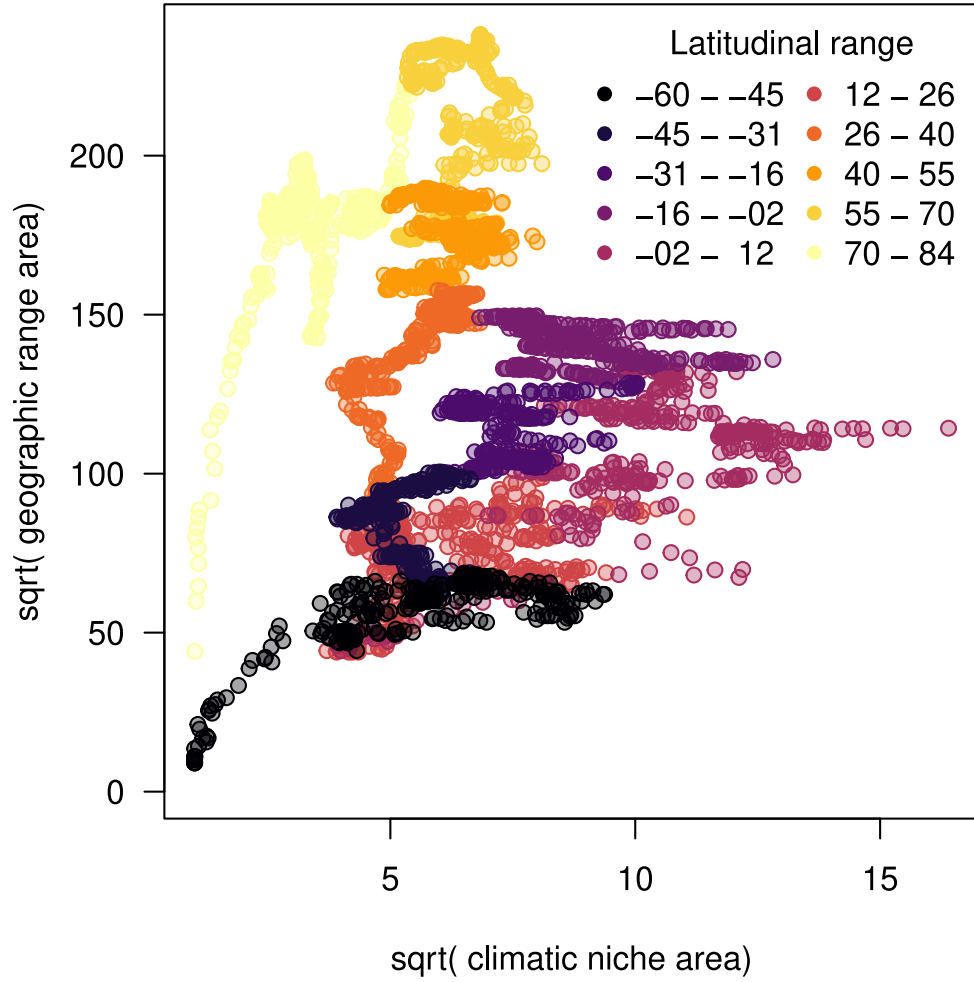


Figure 4: The potential relationship between geographic range size and climatic niche area is constrained by the amount of available land to colonize and the amount of available niche space. Here, each point is a latitudinal band, corresponding to the amount of land area and climatic niche area in that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

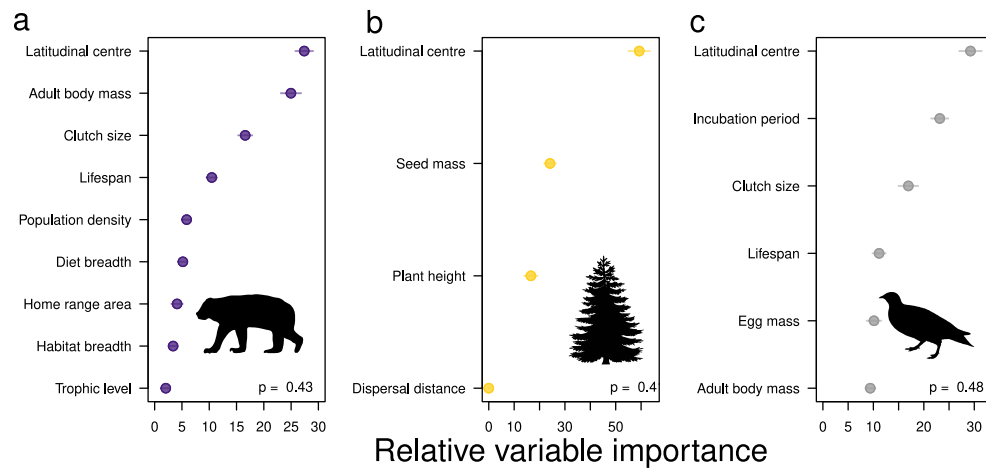


Figure 5: Relative importance values (mean and standard deviation) obtained from boosted regression tree models, which relate residual variation from geographic range size – climatic niche area relationships for mammals, trees, and birds to species traits and latitudinal centre. The importance of species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