

Beyond range size: drivers of species' geographic range structure in European plants

Anna M. Csergő^{1,2*}, Olivier Broennimann^{3,4*}, and Antoine Guisan^{3,4**}, Yvonne M. Buckley^{1,5**}

¹School of Natural Sciences, Zoology, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland (corresponding author)

²Department of Botany and Soroksár Botanical Garden, Szent István University, 1118 Budapest, Hungary

³Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

⁴Institute of Earth Surface Dynamics, University of Lausanne, 1015 Lausanne, Switzerland

⁵ARC Centre of Excellence for Environmental Decisions, School of Biological Sciences, The University of Queensland, Queensland 4072, Australia

* The first authors contributed equally to this work

** The senior authors contributed equally to this work

1 **Beyond range size: drivers of species' geographic range structure in European plants**

2

3 Running title: Geographic range structure of European plant species

4

5 **Abstract**

6 **Aim:** To assess if and how species' range size relates to range structure, if the observed

7 geographic range properties can be retrieved from predicted maps based on species distribution

8 modeling, and whether range properties are predictable from biogeophysical factors.

9 **Location:** Europe

10 **Time period:** Current

11 **Major taxa studied:** 813 vascular plant species endemic to Europe

12 **Methods:** We quantified the size and spatial structure of species' geographic ranges and

13 compared ranges currently occupied with those predicted by species distribution models (SDMs).

14 SDMs were constructed using complete occurrence data from the Atlas Florae Europaeae and

15 climatic, soil and topographic predictors. We used landscape metrics to characterize range size,

16 range division and patch shape structure, and analysed the phylogenetic, geographic and

17 ecological drivers of species' range size and structure using phylogenetic generalized least

18 squares (pGLS).

19 **Results:** Range structure metrics were mostly decoupled from species' range size. We found

20 large differences in range metrics between observed and predicted ranges, in particular for

21 species with intermediate observed range size and occupied area, and species with low and high

22 observed patch size distribution, geographic range filling, patch shape complexity and

23 geographic range fractality. Elevation heterogeneity, proximity to continental coasts, Southerly

24 or Easterly geographic range positions and narrow ecological niche breadth constrained species'

25 observed range size and range structure to different extents. The strength and direction of the
26 relationships differed between observed and predicted ranges.

27 **Main conclusions:** Several range structure metrics, in addition to range size, are needed to
28 adequately describe and understand species' ranges. Species' range structure can be well
29 explained by geophysical factors and species niche width, albeit not consistently for observed
30 and predicted ranges. As range structure can have important ecological and evolutionary
31 consequences, we highlight the need to develop better predictive models of range structure than
32 provided by current SDMs, and we identify the kinds of species for which this is most necessary.

33

34 **Keywords**

35 Connectivity, Ecological Niche Model, Europe, Endemic plant species, Fragmentation,
36 Geographic range size, Habitat suitability, Species' area, Species Distribution Model, Landscape
37 Metrics.

38 **Introduction**

39 Understanding what drives observed large-scale distribution patterns is key to predicting
40 species' future geographic ranges and vulnerability to global change (Broennimann et al. 2006).
41 Analyses of species' geographic ranges have mostly focused on determinants of range size
42 (Morueta-Holme et al. 2013, Sheth et al. 2020). However, examining occupancy patterns
43 throughout a species' range may enable a much better understanding of dispersal and extinction
44 dynamics in response to global environmental changes (Crooks et al. 2011, Cianfrani et al.
45 2018). While occupancy patterns have been extensively used in landscape-level distribution
46 analyses (Wintle et al. 2018), they have only recently and incompletely been applied to species'
47 whole ranges (Brown et al. 1996, Crooks et al. 2011, Pearson et al. 2014, Pagel et al. 2019).

48 Complex interactions and feedbacks between geophysical, ecological and evolutionary
49 processes determine how occupied areas within species' ranges are structured and connected
50 (Gaston 2003, Martinez-Meyer et al. 2012). A species' internal range structure is a consequence
51 of the availability and the spatial arrangement of sites that fulfil the requirements of a species'
52 fundamental niche (Hutchinson's duality; Colwell and Rangel 2009) together with dispersal and
53 persistence dynamics and biotic interactions (Soberon 2007; Guisan et al. 2017, Pagel et al.
54 2019). In turn, constraints due to range structure can modulate a whole suite of ecological and
55 evolutionary processes, such as the routes of continental-scale migrations driven by climatic
56 changes (Hampe and Jump 2011), (meta-)population colonization-extinction dynamics (Manna
57 2017), variation in abundance and genetic diversity within the range (Holt et al. 2002, Pironon et
58 al. 2016), island effects (Covas 2012), or the Allee effect (Lamont et al. 2003). Ultimately,
59 species' internal range structure can foster divergence among populations and provide
60 opportunities for local adaptation, which might in turn alter species' fundamental niche and its

61 spatial expression (Colwell & Rangel 2009). Despite the importance of range structure for
62 ecological and evolutionary processes, quantitative analyses of coarse-scale range changes still
63 largely ignore, or insufficiently account for, this property of species' ranges (but see Crooks et al
64 2011, Pearson et al. 2014, Cianfrani et al. 2018).

65 Quantifying geographic range structure requires unbiased knowledge of species
66 distributions (i.e. consistent throughout the range), which is currently unavailable for most
67 organisms. However, the growing biodiversity repositories and species distribution modeling
68 techniques (SDMs, also called ecological niche or habitat suitability models; Guisan et al. 2017)
69 can provide geographically consistent predictions of species' range-wide distributions. In
70 addition, SDMs can be used to predict not only current but also future range structures, enabling
71 us to move beyond classical range size analyses of climate change impacts (Williams et al. 2008,
72 Cianfrani et al. 2018). A limitation of the approach, however, is that the occupancy of potentially
73 suitable ranges is often incomplete for some species, depending on range shift histories, human
74 land use, dispersal limitations, metapopulation dynamics and species' other biological properties
75 (Svenning et al. 2006, Estrada et al. 2015, Miller and McGill 2017), leading to mismatches
76 between observed and SDM-predicted geographic patterns. Surprisingly, the ability of SDMs to
77 reproduce species' observed internal range geometry has not been thoroughly tested to date. One
78 reason for such limitation might come from the lack of comprehensive distribution data at a large
79 scale to develop models and test predictions. The Atlas Florae Europaeae (AFE) is one of the
80 few datasets with coarse-grained, yet reasonably non-biased presence-absence data across the
81 entire range of a large number of plant species in Europe. Here, we developed SDMs for over
82 800 endemic vascular plant species in AFE and used observed and predicted distributions to
83 quantify species' geographic range size and structure and explore their potential intrinsic and

84 extrinsic determinants. Specifically, across all species, we addressed the following questions: 1)
85 Is range size coupled with range structure? 2) Are measures of range size and structure congruent
86 between observed and predicted ranges? 3) What niche - related and extrinsic biogeophysical
87 factors explain range size and range structure, and are they the same for measures based on
88 observations and predictions?

89

90 **Methods**

91 *Species selection*

92 We accessed species' occurrence data at 50×50 km² resolution from the database of Atlas Florae
93 Europaeae (AFE) which contains 4 473 taxa from 78 vascular plant families. We subset the
94 database to include European endemic species using EvaplantE, the Database on Endemic
95 Vascular Plants in Europe (Hobohm 2014), which we complemented with our own search for
96 species considered endemic by Flora Europaea using the online search engine of Flora Europaea,
97 with the visual aid of biogeographic atlases (Meusel and Jäger, 1992, digitized:
98 <http://chorologie.biologie.uni-halle.de//choro/>). This preselection resulted in 1401 taxa. To
99 decrease the risk of bias in SDM predictions from low number of occurrences (Wisz et al. 2008),
100 we ran a second selection to include only taxa with at least 10 occurrence data points, resulting in
101 864 taxa. We removed 14 species which were unavailable in the phylogenetic tree, and further
102 37 species because the shape of the minimum convex polygon impeded the calculation of
103 particular range metrics, resulting in a final dataset of 813 species. The taxonomy and
104 nomenclature follows the database of the AFE.

105

106 *Species distribution models (range modelling)*

107 To characterise environmental and geographic variation throughout species' ranges, we selected
108 eight climate variables from the CliMond Archive (Kriticos et al. 2014), five soil variables from
109 the International Geosphere-Biosphere Programme IGBP (Global Soil Data Task 2014), and one
110 topographic variable calculated from elevation maps (Jarvis et al. 2008), totaling 14 final
111 predictors: mean diurnal temperature range, temperature annual range, mean temperature of the
112 wettest quarter, mean temperature of driest quarter, mean temperature of warmest quarter, annual
113 precipitation, precipitation seasonality, precipitation of driest quarter, nitrogen density
114 (minimum), nitrogen density (maximum), thermal capacity, wilting point (minimum), wilting
115 point (maximum), slope (selection details in Appendix S1). We built Ensemble of Small Models
116 (ESM; Breiner et al. 2015) and classical species distribution models (SDM; Thuiller et al. 2009)
117 to predict habitat suitability for each species using the selected environmental variables
118 (Appendix S1, Dryad Digital Repository).

119

120 *Range size and range structure metrics*

121 We calculated two range size metrics: geographic range size and occupied area, and four
122 geographic range structure metrics commonly used in landscape ecology, which we adapted to
123 describe how divided and how irregular geographic ranges are. Range structure metrics included
124 two range division metrics (patch size distribution, geographic range filling) and two patch shape
125 metrics (patch shape complexity, geographic range fractality (Box1, Fig. 1, Appendix S2). These
126 metrics are implemented in FRAGSTATS for ArcGIS (McGarigal et al. 2012) and the
127 *SDMTools* package for R (VanDerWal et al. 2014). We calculated all metrics from maps of
128 observed occurrence and habitat suitability separately, with slight modifications of the original
129 functions (Appendix S2.1). We analysed correlations between different metrics to test for their

130 ability to offer complementary information (Appendix S2.2). We tested the sensitivity of the
131 geographic range size and structure metrics to methodological choices taken to define observed
132 and predicted distributions: **1)** map resolution (raster grain size), **2)** rasterization approaches of
133 the AFE atlas grid, **3)** definition of species' geographic range boundaries, **4)** species distribution
134 modeling technique, **5)** prediction error of species distribution models, **6)** habitat quality
135 thresholds used to define a species' potential distribution (Appendices S2.3 to S2.8).

136

137 *Geographic and ecological drivers of range size and structure*

138 We derived geophysical variables and niche breadth that summarise the geographic and
139 ecological properties of species' ranges. We defined the position of species' geographic range in
140 Europe as the median latitude and longitude of all 50×50 km² grids cells occupied (or predicted
141 suitable), and species' range-wide elevational heterogeneity as the standard deviation of the
142 elevation above sea level across occupied (or predicted suitable) grid cells. We calculated range
143 distance to sea coastline as average distance of each occupied (or predicted suitable) grid cell to
144 the closest shore in Europe, using geographic coordinates in the WGS84 projection map to
145 obtain standardized distances. To calculate niche breadth, we reduced the dimensionality of the
146 climate space (the same climate variables as in the SDM) for all grid cells in Europe to two axes
147 using a PCA analysis (Appendix S3), and we summed the standard deviation of values
148 corresponding to species' occurrences on each of the axes, weighted by the corresponding axis
149 Eigenvalue.

150

151 *Statistical analyses*

152 We first analyzed the relationship between the observed range metrics using Pearson's r and
153 Spearman's Rho correlations. Then, to compare species' range size and structure quantified from
154 observed occurrences and predicted habitat suitability maps respectively, we used
155 phylogenetically uncorrected and corrected paired t-tests (Revell, 2012), using the tree available
156 in Zanne et al. (2014). If any difference between observed and predicted range size and structure
157 metrics was removed upon the phylogenetic correction, that indicated influence of evolutionary
158 history on these differences.

159 To further test whether the magnitude and direction of differences between observed and
160 predicted suitable range metrics were predictable from the observed range metrics or
161 phylogenetically structured, we fitted Phylogenetic Generalized Least Squares models (pGLS;
162 Freckleton et al. 2002) of the log range metric (m) response ratio for predicted (p) and occupied
163 (o) ranges, with the range metric calculated from observed occurrences (m_o) as the explanatory
164 variable as main and quadratic effects: $\ln\left(\frac{m_p}{m_o}\right) \sim m_o + m_o^2$. If observed and predicted metrics
165 provide the same information, the response should be constant and set to zero (i.e. $\ln(1)$). By
166 extension, these models indicate for what kinds of observed ranges inferring observed from
167 predicted ranges is the most problematic.

168 Finally, to test how geographic position, elevation heterogeneity and proximity to seacoast of
169 species' ranges, as well as species' niche breadth affected the different range properties, we fitted
170 pGLS models to each of the six observed and predicted metrics separately. The pGLS models
171 were fitted to include the interactions between niche breadth and mean range latitude and
172 between niche breadth and mean range longitude, and the main effects of elevational
173 heterogeneity and proximity to seacoast: $m \sim \textit{niche breadth} \times \textit{latitude} +$
174 $\textit{niche breadth} \times \textit{longitude} + \textit{SD elevation} + \textit{distance to seacoast}$. To make effect sizes

175 comparable across different metrics, all explanatory variables were centered on zero with unit
176 variance. To meet model assumptions, geographic range size, occupied (or suitable) area and
177 patch size distribution were log-transformed in all analyses. The phylogenetic t-tests were fitted
178 with the ‘phytools’ package (Revell 2012) and the pGLS models with the ‘caper’ package (Orme
179 et al. 2013) in R 3.2.3 (R Core Development Team 2015).

180

181 **Results**

182 *Properties and correlations of observed range metrics*

183 Geographic range metrics quantified from observed distributions were highly variable between
184 species (Fig. 2). As expected, the range size metrics, i.e. geographic range size and occupied
185 area, were highly positively correlated (Pearson’s $r = 0.91$). Patch size distribution was positively
186 correlated with occupied area (Pearson’s $r = 0.682$) but only weakly positively correlated with
187 geographic range size (Pearson’s $r = 0.34$), i.e. large occupied patches added up to larger
188 occupied areas but not necessarily to larger range sizes, and large range sizes were not
189 necessarily composed of large patches. Geographic range filling was negatively correlated with
190 geographic range size (Spearman’s $Rho = -0.627$), indicating that, in contrast to species with
191 small geographic ranges, species with large geographic ranges had extended unoccupied areas
192 within the species’ range boundaries. Correlations between range structure metrics were
193 generally weak (Pearson’s r or Spearman’s $Rho < 0.351$), highlighting their complementarity in
194 describing different facets of range structure. An exception was patch size distribution, which
195 was more strongly, positively correlated with geographic range filling, suggesting that species
196 establishing large patches are more likely to fill out species’ ranges. Patch size distribution was

197 also positively correlated with patch shape complexity, indicating spatial constraints on the
198 geometrical complexity of small sized patches (Spearman's Rho > 0.512) (Appendix S2.2).
199
200 The geographic range metrics were not sensitive to map resolution, rasterization approach of
201 occurrence data, or approaches taken to estimate species' geographic range boundaries
202 (Appendix S2.3.1- S2.3.3). For a large number of species, classical SDM projections tended to
203 underestimate geographic range size, occupied area and geographic range filling compared to the
204 ESM projections, and the patch shape metrics calculated from maps produced with the two
205 different techniques were poorly correlated (Appendix S2.3.4). Geographic range metrics based
206 on binary maps produced by different thresholding methods were highly correlated, with few
207 exceptions (Appendix S2.3.5). All results presented in the main text are thus based on binary
208 maps thresholded using MaxTSS. The area metrics and the range division metrics decreased, and
209 the patch shape metrics increased significantly with increasing habitat suitability threshold
210 ($p < 0.05$) (Appendix S2.3.6), indicating lower availability and higher irregularity of highly
211 suitable habitat patches within species' ranges.

212

213 *Differences between observed and predicted range metrics*

214 The phylogenetically uncorrected paired t-tests indicated that mean geographic range size,
215 occupied area, patch size distribution and geographic range filling of observed ranges were
216 significantly lower than those based on predicted suitable ranges ($|t| > 5.90$, $p < 0.001$). In
217 contrast, the phylogenetically corrected paired t-test indicated that only geographic range filling
218 was significantly different, and lower, for observed compared to predicted ranges ($|t| = 5.93$, $p <$
219 0.001) (Fig. 3, Appendix S4.1). Accordingly, tests of phylogenetic signal in each of the range

220 structure metrics separately indicated significant evolutionary relationships between species in
221 observed and predicted geographic range size, occupied area and patch size distribution (Pagel's
222 $\lambda > 0.731$, $p < 0.001$) (Appendix S4.2). Therefore, differences between observed and predicted
223 values for these three metrics were likely structured by evolutionary relatedness between species.

224

225 The log response ratio between observed and predicted metrics differed in magnitude across
226 metrics (Fig. 4, Fig. 5, Appendix S4.3). Median log response ratios closest to zero with the
227 narrowest range, and hence best observed to predicted suitable matches, were for the patch shape
228 metrics (Fig. 4e, f) and the largest medians and ranges of log response ratios, and hence largest
229 divergence between observed and predicted suitable, were detected for the range size metrics and
230 patch size distribution (Fig. 4a, b, c). There were major differences between range size and range
231 structure metrics in the shape of the log response ratio relative to the observed metric. In the
232 range size metrics, higher and largely positive log response ratios were common for medium
233 sized observed values (positive quadratic effects, $p < 0.001$) (Fig. 4a, b). For the range structure
234 metrics, log response ratios were often higher for low observed values and, conversely, lower for
235 high observed values (negative linear or quadratic effects, $p < 0.001$) (Figure 4c - 4f). Exceptions
236 were extreme high values of patch size distribution and geographic range filling, which showed
237 higher, or no change in log response ratio respectively. With the exception of geographic range
238 fractality ($R_{sq} = 0.75$), models of all other metrics explained a moderate proportion of variance
239 in data ($R_{sq} < 0.48$).

240

241 *Congruent and divergent drivers of the observed and predicted range metrics*

242 Congruent effects of biophysical and niche variables on observed and predicted range metrics
243 were mostly found for the observed and predicted range size metrics, which were significantly
244 positively correlated with species' niche breadth and species' median range latitude. In addition,
245 there was a small significant negative interaction between species' niche breadth and species'
246 latitude for both predicted and observed range size metrics. Thus, species' observed and
247 predicted range size and area increased more rapidly with increasing niche breadth in species of
248 Southern Europe, compared to species of Northern Europe (Fig. 6a, e, Fig. 7a, e, Appendix S4.4,
249 S4.5). Further, both the occupied and suitable area were significantly smaller in species in
250 topographically heterogeneous areas, and both the occupied and suitable area and range size
251 were significantly smaller in species distributed closer to Europe's coastline (Fig. 6c, d, Fig. 7c,
252 d, Appendix S4.4, S4.5).

253 Divergent effects of biophysical and niche variables on observed and predicted range size
254 metrics were found only in two instances: 1) occupied area was significantly lower in species of
255 Eastern Europe, but suitable area was not correlated with median range longitude (Fig. 6b, Fig.
256 7b, Appendix S4.4, S4.5), and 2) observed range size was significantly lower in species
257 distributed over large elevation gradients, but predicted range size was not related to topographic
258 complexity (Fig. 6c, 7c, Appendix S4.4, S4.5).

259 Of the range structure metrics, congruent effects were only found for observed and predicted
260 range filling, both significantly negatively correlated with median range latitude and species'
261 niche breadth, and there was a significant positive interaction of niche breadth and latitude on
262 this metric (Fig. 6a, e, Fig. 7a, e, Appendix S4.4, S4.5). Thus, range filling decreased strongly
263 with increasing niche breadth in Southern Europe, but to a lesser extent in Northern Europe.

264 Divergent effects of geographic and ecological variables were common between observed and
265 predicted range structure metrics. While for predicted patch size distribution we found a
266 significant negative interaction between niche breadth and latitude, the observed patch size
267 distribution was affected, positively, only by species' niche breadth (Fig. 6e, Fig. 7a, e,
268 Appendix S4.4, S4.5). Predicted patch size distribution was not correlated with range longitude,
269 but the observed patch size distribution was significantly lower in species of Eastern Europe
270 (Fig. 6b, Fig. 7b, Appendix S4.4, S4.5). Predicted patch size distribution decreased significantly
271 with increasing topographic heterogeneity and nearer to Europe's coastlines, but the observed
272 patch size distribution was not affected significantly by either parameter (Fig. 7c, d, Appendix
273 S4.4, S4.5).

274 Predicted geographic range filling was significantly lower in topographically complex areas, but
275 the observed range filling did not vary with topographic heterogeneity (Fig. 7c, Appendix S4.4,
276 S4.5).

277 Predicted patch shape complexity was significantly higher in species distributed at high latitudes,
278 longitudes and significantly lower near coastlines, but the observed metric did not respond to any
279 of these parameters. Predicted patch shape complexity decreased significantly with increasing
280 topographic heterogeneity, but the opposite effect was found in the observed metric (Fig. 6c, 7c,
281 Appendix S4.4, S4.5).

282 Predicted geographic range fractality was significantly lower at high latitudes, but the observed
283 metrics did not vary with latitude. Predicted range fractality did not correlate with longitude, but
284 there was a significant negative interaction between niche breadth and longitude for the observed
285 geographic range fractality, hence in species with larger niche breadth, the observed complexity
286 of patch shapes was more even across a range of patch sizes than predicted in species of Eastern

287 Europe (Fig. 7a, b, d, Appendix S4.4, S4.5). Predicted geographic range fractality increased
288 significantly with increasing topographic heterogeneity, but the opposite effect was found in the
289 observed metric (Fig. 6c, 7c, Appendix S4.4, S4.5).

290 The pGLS models explained a large proportion of variation in observed and predicted
291 geographic range size, occupied (suitable) area and predicted patch size distribution ($0.669 < \text{adj.}$
292 $R^2 < 0.836$), and a lower proportion of variation in observed patch size distribution and
293 observed and predicted geographic range filling, patch shape complexity and geographic range
294 fractality ($\text{adj. } R^2 \text{ observed} < 0.328$ and $\text{predicted} < 0.186$) (Appendix S4.4).

295 Phylogenetic signal was detected only in pGLS models of occupied area (Model λ : 0.222,
296 Appendix S4.4).

297

298 **Discussion**

299 Using complete occurrence data from atlas maps at ~50 km resolution and species distribution
300 models for over 800 plant taxa endemic to Europe, we show that the internal structure of species'
301 observed ranges varies substantially among species, and cannot be inferred consistently from
302 measures of range size, nor from SDM predictions. Differences between species' observed and
303 predicted range and area size and patch size distribution are phylogenetically structured. While
304 geographical factors and niche breadth affect the size and structure of species' ranges, the
305 explanatory variables differ qualitatively and quantitatively between observed and predicted
306 range metrics. Range structure metrics are important to consider independently of, or in addition
307 to, range size metrics when interpreting biogeographic patterns of observed or predicted
308 occupancy.

309

310 *Range structure is largely decoupled from range size*

311 Range size metrics, in particular those that incorporate both unoccupied and occupied areas
312 within the range boundaries are not congruent with species' internal range structure. The
313 negative correlation between range size and range filling supports the view that mechanisms
314 shaping species' range size (extent of occurrence) and the ability of species to fully occupy it
315 (range occupancy) may differ substantially (Gaston et al. 2003). The observed patch size
316 distribution was better correlated with occupied area than with the observed geographic range
317 size, suggesting that species' occupied area could be linked to processes that create abundance-
318 occupancy patterns at the landscape scale (Freckleton et al. 2006, Meloni et al. 2017), while
319 species' range size can additionally be shaped by spatio-temporal processes that generate
320 occurrence patterns, such as long-distance dispersal or spatial rearrangement of populations with

321 geographic range shifts (Hampe and Jump 2011). Analysing which landscape-level processes
322 scale up to structure species' biogeographic ranges has remained largely unexplored (Kent
323 2007), yet our results suggest that it may be key to understanding and predicting internal range
324 rearrangements in changing environments.

325

326 *Differences between observations and predictions are not consistent among and within range*
327 *size and structure metrics*

328 Differences between observed and predicted geographic range size, area and patch size
329 distribution were due to phylogenetic relationships between species. One reason for the
330 phylogenetically structured mismatches between these predicted and observed metrics could be a
331 higher range disequilibrium with the environment of particular groups of closely related species,
332 which would lower the ability of SDMs to correctly predict distributions (Guisan et al. 2017).
333 Environmental stability may promote phylogenetic conservatism and it also modulates species'
334 range size (Morueta-Holme et al. 2013, Zacaï et al. 2017). We hypothesize that range
335 disequilibrium due to large-scale climatic events on the continent (e.g. glaciations) or human
336 land use limiting dispersal (Miller and McGill 2017) were more dramatic in selected groups of
337 phylogenetically related species, while species with closer observed to predicted matches
338 benefited from more stable climates. Another hypothesis posits that phylogenetically older
339 species may have more fragmented ranges than their younger relatives (Sheth et al. 2020), due to
340 e.g., increasing complexity of biotic interactions with species' age. While these hypotheses need
341 further testing, our results underscore that we lack a good understanding of the phylogenetic
342 structuring of species' ranges (Gaston et al. 2003, Sheth et al. 2020).

343 The only metric with higher predicted relative to observed values, independently of the
344 phylogenetic relationships between species, was geographic range filling. Consequently, it is
345 very common that SDMs extend species' range boundaries with remote suitable habitat patches
346 that are likely to fall beyond species' normally observed colonisation abilities (Estrada et al.
347 2015).

348

349 Different sets of species were responsible for the divergence between observed and predicted
350 range size and range structure metrics. For the range size metrics, inferring realized from
351 predicted range and area sizes was the most problematic for species with intermediate observed
352 values of these metrics. This may arise due to both methodological and ecological constraints:
353 for small ranges, the convergence between observations and predictions may be a modeling
354 artefact due to models being more easily overfit; for large ranges, the convergence might be due
355 to geographical constraints from landmass availability and continental barriers, which can drive
356 species' global range size especially in highly dispersive species (Broennimann et al. 2006,
357 Sheth et al. 2020). Alternatively, both narrow climate endemics of Europe (Ohlemüller et al.
358 2008) and species with large niches might be less prone to modeling biases from niche
359 truncation.

360 For the range structure metrics, inferring realized from predicted patch size distribution and
361 range filling was problematic in species with the smallest observed values of these metrics. In
362 such species with constrained range occupancy - but not necessarily small range sizes -,
363 overpredictions from SDMs due to missing limiting variables (dispersal limitation, biotic
364 interactions, human impact, etc.) are highly likely. Consequently, the predicted patch size
365 distributions and spatial arrangement of suitable habitat patches within species' ranges are not

366 realized, or conversely, are eroded, and theoretical inferences about the associated spatial
367 processes need careful testing.
368 Discrepancies between predicted and observed patch shapes raises modeling difficulties of a
369 different nature. The predicted simplification of the complex observed shapes, which as a result
370 modulates discrepancies between predicted and observed range fractality, suggests that SDMs
371 are blurring patch boundaries (also indicated by the lower variance of predicted than of the
372 observed patch shape metrics in Fig. 3). Distribution borders theoretically self-arrange along
373 complex geometrical patterns (Oborny 2018) that are very likely lost during the predictive
374 modelling process and over coarse spatial resolutions.
375 These important divergences between metrics and between species severely complicate current
376 and future predictions of species' internal range rearrangement under climate change, and we
377 cannot hereafter ignore the direction and magnitude of the differences when we infer realized
378 distributions from predicted habitat suitability maps.

379

380 *Divergent responses of observed and predicted range size and structure to geographical and*
381 *ecological factors*

382 Species' niche breadth was a strong biogeographic determinant of species' ranges, with similar
383 effects on the predicted and observed range metrics. The widely recognized niche effects on
384 range and area size (Moore et al. 2018, Sheth et al. 2020) were associated with further niche
385 effects on species' range structure (observed and predicted geographic range filling and patch
386 size distribution, and predicted range fractality). Niche breadth can therefore be a useful tool in
387 global predictions of species' range structure and internal range rearrangements, especially in a
388 climate change context (Broennimann et al. 2006, Connor et al. 2018). However, niche effects on

389 range structure metrics were modulated by range latitude and longitude, and the relatively
390 homogeneous climate over large areas in Northern Europe (Ohlemüller et al. 2008) might have
391 enabled species with different niche breadths, that include the Northern European climate, to
392 acquire relatively large range and area sizes, filled with large suitable patches. This result
393 contradicts the view that towards higher latitudes larger range sizes are due solely to species'
394 increasing ecological tolerance (Rapoport's rule:, Rapoport 1982, Stevens 1989). Ecological
395 tolerance does not seem to exclusively govern geographic ranges (Gaston and Chown 2016,
396 Payne and Smith 2017, Moore et al. 2018), instead it might modulate species' range size and
397 structure within the spatial availability of geophysical conditions, according to the principles of
398 Hutchinson's niche-biotop duality (Colwell and Rangel 2009).

399 Constraints on species' predicted and observed area and range sizes in Southern Europe, in
400 topographically heterogeneous areas, as well as near the coastlines were often, but not always,
401 associated with additional constraints on range structure. The historical imprints of the higher
402 climate stability in Southern Europe during glaciations, associated with a diversity of small
403 climate pockets (Ohlemüller et al. 2008), the higher rates of speciation in topographically
404 complex areas (Morueta-Holme et al. 2013, Steinbauer et al. 2016), and the geographical barriers
405 imposed by the continental coastlines (Broennimann et al. 2006) on species ranges could
406 therefore also be manifest in species' internal range geometry, with potentially important
407 ecological and evolutionary consequences that need further exploration.

408 However, most effects of geographical parameters on the predicted range metrics were not
409 transferable to the observed metrics, and we suspect biological and historical, as well as
410 methodological reasons underlying these inconsistencies. The relaxed dependency of predicted
411 patch size distribution on niche breadth towards the North, undetected in observed patch size

412 distributions, could be due to at least two factors: 1) patch size distribution is modulated by
413 species' biological properties that manifest regardless of species' range position, such as
414 dispersal ability (Freckleton et al. 2006, Pagel et al. 2019) and 2) there is a strong postglacial
415 dispersal lag in species distributed in Northern Europe, which leaves large parts of suitable
416 habitat patches unoccupied (Svenning et al. 2006). Historical factors or anthropic pressures
417 might have similarly reshuffled, or eroded, theoretical geographical patterns of Eastern Europe
418 (Kajtoch et al. 2016), where we revealed an observed, but not predicted, decline in area and
419 patch size distribution, as well as predicted, but not observed, increase in patch shape
420 complexity. Topography and geographical barriers were not enough to explain realized patch
421 size patterns either, because small observed patch size distributions due to small average patch
422 size or small overall area were not only frequent in highlands and near the coastlines as expected,
423 but also in the more homogenous lowlands and inland. Thus, human land use, colonization and
424 fragmentation histories, or species' intrinsic properties, are most likely modifying the predicted
425 theoretical internal range geometries, obscuring the accuracy of future range forecasts (Svenning
426 et al. 2006, Estrada et al. 2014, Miller and McGill 2017).

427

428 *Limitations and future directions*

429 Due to the coarse resolution of our dataset, we only tested drivers of range size and structure that
430 may operate over large biogeographic scales, such as niche breadth, phylogeny, coarse scale
431 geomorphology, large geographic barriers and coarse geographic gradients. The repertoire of
432 metrics that describe species' range size and structure structure can definitely be expanded in the
433 future and approaches can certainly be refined (Sheth et al. 2020). Some of the metrics chosen in
434 this study were weakly responsive to the chosen geophysical and ecological drivers but may better

435 respond to other factors yet to be tested. Smaller effect sizes with large confidence intervals for
436 predictions of patch shape complexity and range fractality metrics emphasised that these metrics
437 are poorer predictors of processes at global scales, similarly to findings of Pearson et al. (2014).
438 Coarse map resolutions can affect global range predictions (Connor et al. 2018), and the smoothing
439 of the range shapes by the distribution models may become problematic when aiming for accurate
440 projections of species' internal range geometry into the future. Further efforts should be made to
441 identify which aspects of species' biology or of the geophysical space are linked to biological
442 patterns at given spatial resolutions (Connor et al. 2018, Pagel et al. 2019), and to improve species
443 distribution models with finer predictors able to better reflect the peculiarities of species' internal
444 range structure.

445

446 **Conclusions**

447 Our results have important implications for the development and interpretation of predictive
448 models of species expansion, persistence and extinction in response to climate change.

449 Fragmentation of occupied area or suitable habitat has already been identified as a better
450 predictor of extinction risk than range size (Pearson et al. 2014, Crooks et al. 2017), and we
451 propose that factors influencing different aspects of species' range structure, in addition to those
452 that influence range size, may be useful conservation targets (as in Cianfrani et al. 2018).

453 The predictable inconsistencies between observed and current predicted range structure metrics
454 and their drivers enable us to better approximate uncertainties in species predictions. Some of the
455 discrepancies might stem from species' inability to occupy all habitats predicted suitable, while
456 others may be linked to our inability to capture accurately species' responses to environmental
457 constraints when modeling distributions. By examining which species are sensitive to different

458 properties of species' ranges, and the sensitivity of populations to changing range configurations,
459 we can refocus conservation efforts to search for critical range structure thresholds for species at
460 the highest risk.

461

462

463 **Acknowledgements**

464 We thank AFE for distribution data. We thank Carsten Hobohm for kindly providing the list of
465 European endemic species in EvaplantE, the database on endemic vascular plants in Europe. We
466 thank Ruben Mateo for help with CGRS grid calculations in early stages of the project and
467 Martha Liliana-Serrano and Liam Revell for readily helping with phylogenetic analyses. We
468 thank Beáta Oborny for very useful comments on manuscript versions in the making. AMC was
469 funded by the Marie Skłodowska-Curie Individual Fellowship GEODEM-658651 under the EU
470 Horizon 2020 Framework Programme for Research and Innovation of the European Research
471 Council. YMB was supported by an Irish Research Council Laureate Award IRCLA/2017/60.
472 AG was supported by a Swiss National Science Foundation grant (CR23I2_162754).

473

474

475 **Biosketch**

476 Anna M. Csergő led this work as a Marie Skłodowska-Curie postdoctoral research fellow at
477 Trinity College Dublin, Ireland and then as associate professor at Szent István University,
478 Hungary (<https://annamariacsergo.weebly.com/>). Together with the co-authors she is interested
479 in how species' life history and population-level processes interact with the geometry and
480 suitability of habitat fragments within a species' range to modulate global persistence patterns.

References

1. Breiner, F.T., Guisan, A., Bergamini, A. & Nobis, M.P. (2015) Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, **6**, 1210–1218.
2. Colwell, R.K. & Rangel, T.F. (2009) Hutchinson’s duality: The once and future niche. *Proceedings of the National Academy of Sciences*, **106**, 19651–19658.
3. Connor, T., Hull, V., Viña, A., Shortridge, A., Tang, Y., Zhang, J., Wang, F. & Liu, J. (2018) Effects of grain size and niche breadth on species distribution modeling. *Ecography*, **41**, 1270–1282.
4. Covas, R. (2012) Evolution of reproductive life histories in island birds worldwide. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 1531–1537.
5. Crooks, K.R., Burdett, C.L., Theobald, D.M., Rondinini, C. & Boitani, L. (2011) Global patterns of fragmentation and connectivity of mammalian carnivore habitat. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2642–2651.
6. Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002) Phylogenetic analysis and comparative data: A test and review of evidence. *American Naturalist*, **160**, 712–726.
7. Freckleton, R.P., Noble, D. & Webb, T.J. (2006) Distributions of habitat suitability and the abundance-occupancy relationship. *The American naturalist*, **167**, 260–275.
8. Gaston, K.J. (2003) *The Structure and Dynamics of Geographic Ranges*, Oxford University Press.
9. Gaston, K.J. & Chown, S.L. (2016) Why Rapoport’s rule does not generalise. *Oikos*, **86**, 584–590.

10. Guisan, A., Thuiller, W. & Zimmermann, N.E. (2017) *Habitat Suitability and Distribution Models*, Cambridge University Press, Cambridge.
11. Hobohm, C. ed. (2014) *Endemism in Vascular Plants*, Springer Netherlands, Dordrecht.
12. Jarvis, A., H.I. Reuter, A. Nelson, E.G. (2008) Hole-filled SRTM for the globe Version: SRTM 90m Database.
13. Kajtoch, Ł., Cieślak, E., Varga, Z., Paul, W., Mazur, M.A., Sramkó, G. & Kubisz, D. (2016) Phylogeographic patterns of steppe species in Eastern Central Europe: a review and the implications for conservation. *Biodiversity and Conservation*, **25**, 2309–2339.
14. Kriticos, D.J., Webber, B.L., Leriche, A., Ota, N., Macadam, I., Bathols, J. & Scott, J.K. (2012) CliMond: Global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution*, **3**, 53–64.
15. McGarigal, K., Cushman, S. & Ene, E. (2012) FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps.
16. Meusel, H. & Jager, E.J. (1992) Vergleichende Chorologie der Zentraleuropaischen Flora (digitized). Volumes I-III.
17. Morueta-Holme, N., Enquist, B.J., McGill, B.J., Boyle, B., Jørgensen, P.M., Ott, J.E., Peet, R.K., Šímová, I., Sloat, L.L., Thiers, B., Violle, C., Wisser, S.K., Dolins, S., Donoghue, J.C., Kraft, N.J.B., Regetz, J., Schildhauer, M., Spencer, N. & Svenning, J.C. (2013) Habitat area and climate stability determine geographical variation in plant species range sizes. *Ecology Letters*, **16**, 1446–1454.
18. Oborny, B. (2018) Scaling laws in the fine-scale structure of range margins. *Mathematics*, **6**, 15–18.

19. Ohlemüller, R., Anderson, B.J., Araújo, M.B., Butchart, S.H.M., Kudrna, O., Ridgely, R.S. & Thomas, C.D. (2008) The coincidence of climatic and species rarity: High risk to small-range species from climate change. *Biology Letters*, **4**, 568–572.
20. Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N. & Will, P. (2013) caper: Comparative Analyses of Phylogenetics and Evolution in R.
21. Payne, N.L. & Smith, J.A. (2017) An alternative explanation for global trends in thermal tolerance. *Ecology Letters*, **20**, 70–77.
22. Pearson, R.G., Stanton, J.C., Shoemaker, K.T., Aiello-Lammens, M.E., Ersts, P.J., Horning, N., Fordham, D.A., Raxworthy, C.J., Ryu, H.Y., Mcnees, J. & Akçakaya, H.R. (2014) Life history and spatial traits predict extinction risk due to climate change. *Nature Climate Change*, **4**, 217–221.
23. Pironon, S., Papuga, G., Villellas, J., Angert, A.L., Garcia, M.B. & Thompson, J.D. (2016) Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biol Rev Camb Philos Soc*.
24. Rapoport, E. (1982) *Areography: geographical strategies of species*, Pergamon Press.
25. Revell, L.J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217–223.
26. Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, **10**, 1115–1123.
27. Steinbauer, M.J., Field, R., Grytnes, J.A., Trigas, P., Ah-Peng, C., Attorre, F., Birks, H.J.B., Borges, P.A.V., Cardoso, P., Chou, C.H., De Sanctis, M., de Sequeira, M.M., Duarte, M.C., Elias, R.B., Fernández-Palacios, J.M., Gabriel, R., Gereau, R.E., Gillespie, R.G., Greimler, J., Harter, D.E.V., Huang, T.J., Irl, S.D.H., Jeanmonod, D., Jentsch, A.,

- Jump, A.S., Kueffer, C., Nogué, S., Otto, R., Price, J., Romeiras, M.M., Strasberg, D., Stuessy, T., Svenning, J.C., Vetaas, O.R. & Beierkuhnlein, C. (2016) Topography-driven isolation, speciation and a global increase of endemism with elevation. *Global Ecology and Biogeography*, **25**, 1097–1107.
28. Svenning, J.C., Normand, S. & Skov, F. (2006) Range filling in European trees. *Journal of Biogeography*, **33**, 2013–2018.
29. VanDerWal, J., Falconi, L., Januchowski, S., Shoo, L., Storlie, C. & Vanderwal, M.J. (2015) SDMTools: Species Distribution Modelling Tools: Tools for processing data associated with species distribution modeling exercises. 46.
30. Williams, S.E., Shoo, L.P., Isaac, J.L., Hoffmann, A.A. & Langham, G. (2008) Towards an integrated framework for assessing the vulnerability of species to climate change. *PLoS Biology*, **6**, e325.
31. Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., Elith, J., Dudík, M., Ferrier, S., Huettmann, F., Leathwick, J.R., Lehmann, A., Lohmann, L., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M.C., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S.E. & Zimmermann, N.E. (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.
32. Zacaï, A., Fara, E., Brayard, A., Laffont, R., Dommergues, J.L. & Meister, C. (2017) Phylogenetic conservatism of species range size is the combined outcome of phylogeny and environmental stability. *Journal of Biogeography*, **44**, 2451–2462.
33. Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., Fitzjohn, R.G., McGlenn, D.J., O’Meara, B.C., Moles, A.T., Reich, P.B., Royer, D.L., Soltis, D.E.,

- Stevens, P.F., Westoby, M., Wright, I.J., Aarssen, L., Bertin, R.I., Calaminus, A., Govaerts, R., Hemmings, F., Leishman, M.R., Oleksyn, J., Soltis, P.S., Swenson, N.G., Warman, L. & Beaulieu, J.M. (2014) Three keys to the radiation of angiosperms into freezing environments. *Nature*, **506**, 89–92.
34. Wintle, B.A., Kujala, H., Whitehead, A., Cameron, A., Veloz, S., Kukkala, A., Moilanen, A., Gordon, A., Lentini, P.E., Cadenhead, N.C.R. & Bekessy, S.A. (2019) Global synthesis of conservation studies reveals the importance of small habitat patches for biodiversity. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 909–914.
35. Miller, K.M. & McGill, B.J. (2017) Land use and life history limit migration capacity of eastern tree species. *Global Ecology and Biogeography*, 1–11.
36. Stevens, G.C. (2016) The Latitudinal Gradient in Geographical Range : How so Many Species Coexist in the Tropics. *The American Naturalist*, **133**, 240–256.
37. Moore, T.E., Bagchi, R., Aiello-Lammens, M.E., Schlichting, C.D. & Timothy Moore, C.E. (2018) Spatial autocorrelation inflates niche breadth-range size relationships. *Global Ecol Biogeogr*, 1–11.
38. Manna, F., Pradel, R., Choquet, R., Fréville, H. & Cheptou, P.O. (2017) Disentangling the role of seed bank and dispersal in plant metapopulation dynamics using patch occupancy surveys. *Ecology*, **98**.
39. Lamont, B.B., Klinkhamer, P.G.L. & Witkowski, E.T.F. (1993) Population fragmentation may reduce fertility to zero in *Banksia goodii*-a demonstration of the Allee effect. *Oecologia*, **94**.

40. Cianfrani, C., Broennimann, O., Loy, A. & Guisan, A. (2018) More than range exposure: Global otter vulnerability to climate change. *Biological Conservation*, **221**, 103–113.
41. Broennimann, O., Thuiller, W., Hughes, G.O., Midgley, G.F., Alkemade, R. & Guisan, A. (2006) Do geographic distribution, niche property and life form explain plants' vulnerability to global change? *Global Change Biology*, **12**, 1079–1093.
42. Sheth, S.N., Morueta-Holme, N. & Angert, A.L. (2020) Determinants of geographic range size in plants. *New Phytologist*, nph.16406.
43. Pagel, J., Treurnicht, M., Bond, W.J., Kraaij, T., Nottebrock, H., Schutte-Vlok, A., Tonnabel, J., Esler, K.J. & Schurr, F.M. (2019) The niche is not the range: Dispersal and persistence shape mismatches between ecological niches and geographic distributions of plants. *bioRxiv*, 526251.
44. Martinez-Meyer, E., Diaz-Porras, D., Peterson, A.T. & Yanez-Arenas, C. (2012) Ecological niche structure and rangewide abundance patterns of species. *Biology Letters*, **9**, 20120637–20120637.
45. Estrada, A., Meireles, C., Morales-Castilla, I., Poschlod, P., Vieites, D., Araújo, M.B. & Early, R. (2015) Species' intrinsic traits inform their range limitations and vulnerability under environmental change. *Global Ecology and Biogeography*, **24**, 849–858.
46. Holt, A.R., Gaston, K.J. & He, F. (2002) Occupancy-abundance relationships and spatial distribution: A review Basic and Applied Ecology. *Basic Appl. Ecol*, **3**, 1–13.
47. Brown, J.H., Stevens, G.C. & Kaufman, D.M. (1996) The geographic range: size, shape, boundaries, and internal structure. *Annu. Rev. Ecol. Evol. Syst*, **27**, 597–623.
48. Hampe, A. & Jump, A.S. (2011) Climate Relicts: Past, Present, Future. *Annual Review of Ecology, Evolution, and Systematics*, **42**, 313–333.

49. ***R Core Development Team. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2015).
50. ***The database of Atlas Florae Europaeae. <https://www.luomus.fi/en/database-atlas-florae-europaeae>
51. ***Global Soil Data Task. Global Soil Data Products CD-ROM Contents (IGBP-DIS). Data Set. (2014) Oak Ridge, Tennessee, USA.

Data accessibility statement

Data associated with this manuscript will be made openly available in the Dryad data repository upon acceptance of the manuscript.

Box 1. Geographic range size and range structure terminology

As with the geographic range size measures (Gaston 1996 TREE, Box 1), no standard methodology exists for measuring the structure of species' geographic ranges. We used the following interpretations:

Geographic Range Size: the area of occupied and non-occupied areas within a species' range boundaries (extent of occurrence, Gaston 1996).

Occupied Area: the area of occupied habitat patches (area of occupancy, Gaston 1996)

Patch Size Distribution: the subdivision of a species' geographic range into habitat patches of different sizes relative to species' geographic range size. Likely reflects processes driving abundance-occupancy patterns (McGarigal et al. 2012, modified).

Geographic Range Filling: proportion of a species' geographic range represented by occupied habitat patches within a species' geographic range boundaries. Some earlier interpretations of range filling may refer to occupancy of suitable habitats, for which we suggest "suitable habitat occupancy" as a more suggestive terminology. The term range filling proposed here captures differences between species' geographic range size and occupied area (McGarigal et al. 2012, modified).

Patch Shape Complexity: the shape complexity of habitat patches averaged across all habitat patches within a species' geographic range (McGarigal et al. 2012, modified).

Geographic Range Fractality: the consistency of the shape of habitat patches across different patch sizes. Range fractality applied to range-level analyses can be viewed as a general patch shape complexity metric, however in our interpretation it indicates a non-additive range shape complexity i.e., quantification of pattern repetition over different spatial scales (McGarigal et al. 2012, modified).

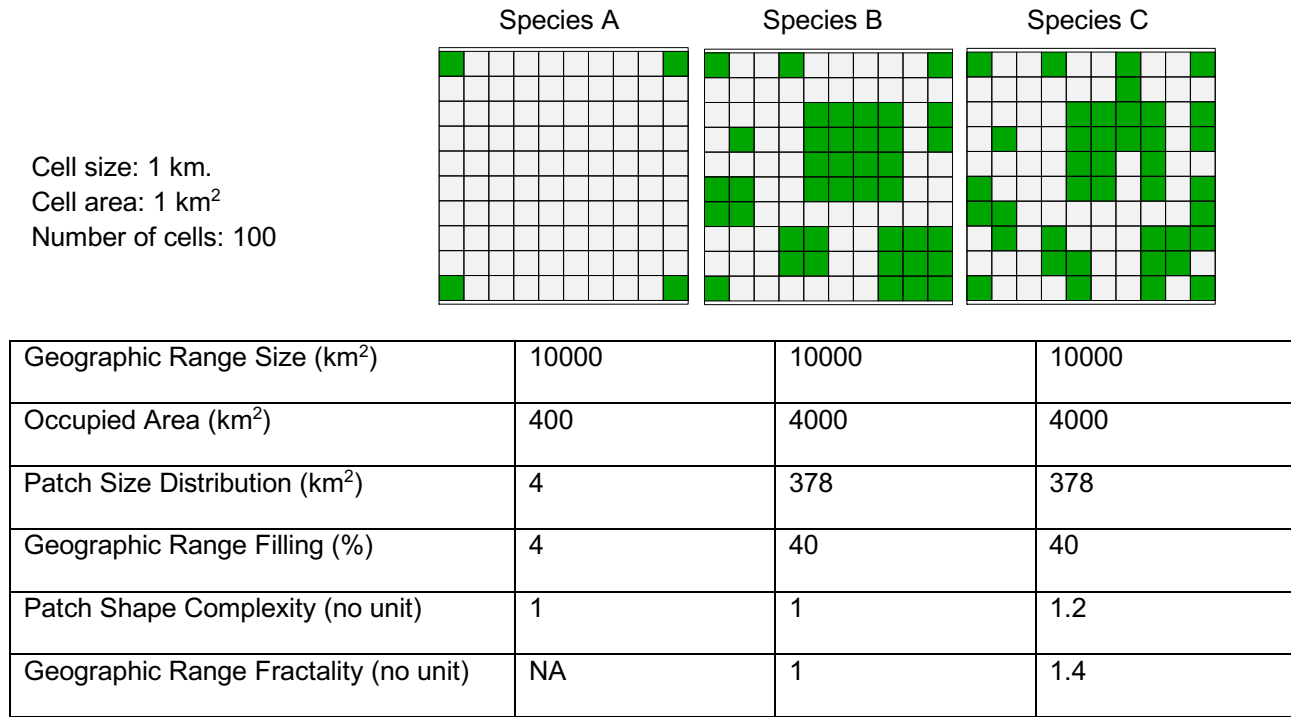


Figure 1. Examples of geographic range size and structure metrics in three theoretical species.

The three species have similar geographic range sizes of ca. 10000 km² but differ in their occupied area and range structure metrics. Species A has low occupied area and fills a small proportion of its range, it also has the lowest patch size distribution. The patch shape complexity for A is low. Species B and C have a high and equal occupied area and fill the range to the same extent, with the same patch size distribution. Species B and C differ in their patch shape complexity and range fractality, with C having on average more complex patch shapes, and the complexity of patches increases with patch size. The geographic range fractality cannot be defined for species A, which has a low number of patches.

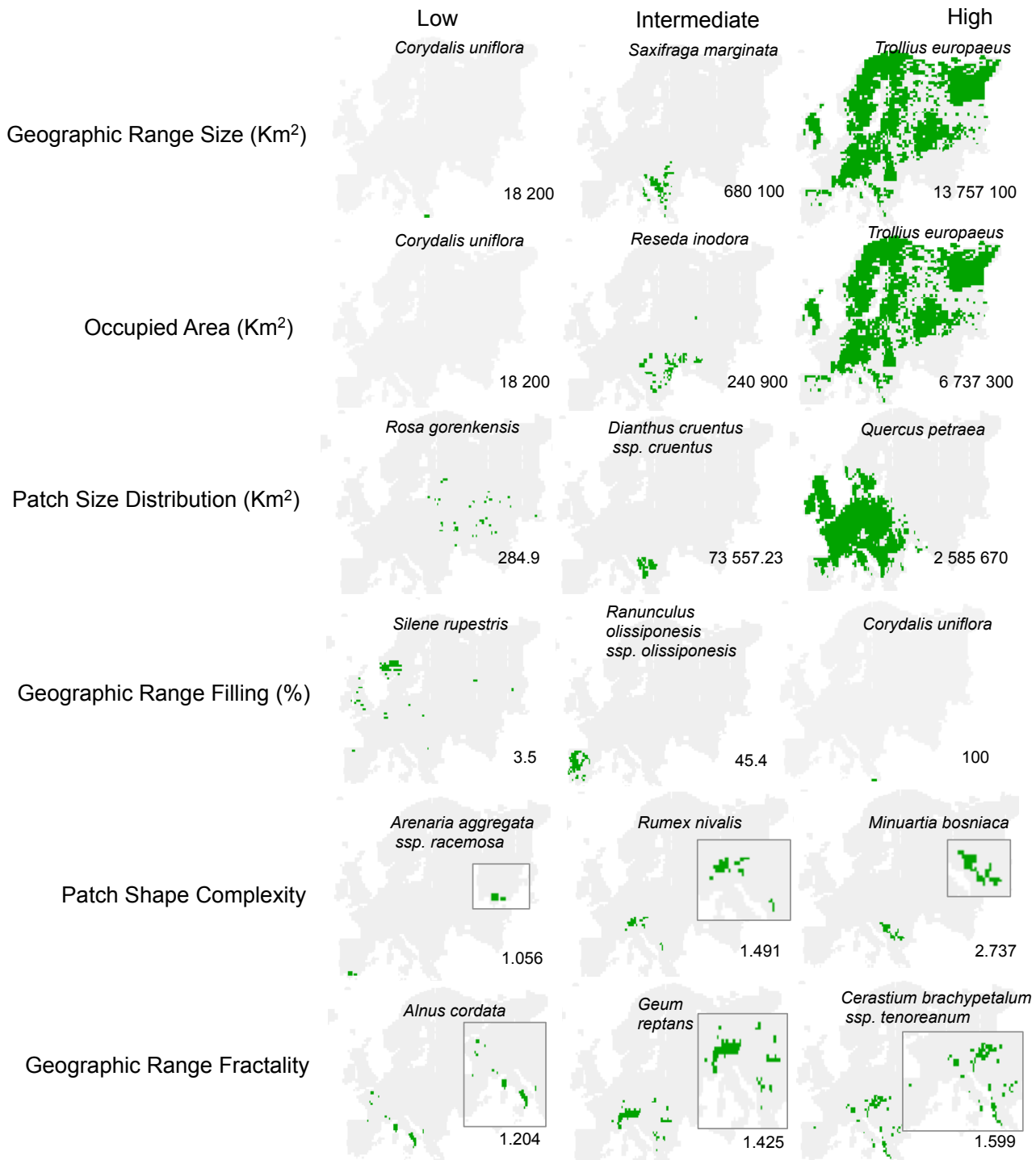


Figure 2. Example species with low, intermediate and high values of geographic range size and structure metrics from a set of 813 European endemic vascular plants, calculated from observed occurrences. Insets represent magnified range maps. Metric values are shown in the bottom right corner of each map.

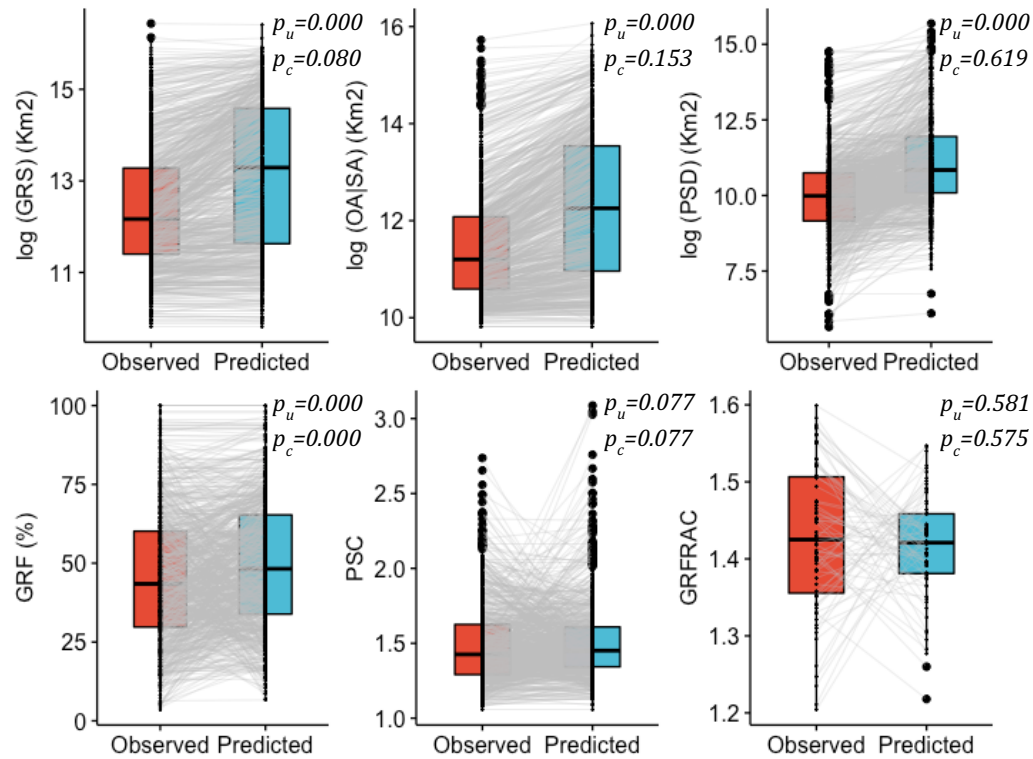


Figure 3. Comparisons of geographic range size and structure metrics quantified from observed and predicted distribution maps for European plant endemics. Significance values of phylogenetically uncorrected (p_u) and corrected (p_c) pairwise t -tests are shown on the upper right corner of each panel. Significant differences between observed and predicted suitable metrics for geographic range size (GRS), occupied area|suitable area (OA|SA) and patch size distribution (PSD) were explained by phylogenetic relatedness. Differences between observed and predicted suitable geographic range filling (GRF) remained significant after phylogenetic correction. No significant differences were found for patch shape complexity (PSC) and geographic range fractality (GRFRAC) in either test. Sample size = 813 species, except for GRFRAC, where analyses were limited to 71 species. The maximum value of the y axis for GRS and OA|SA nears the area of the European continent: $\log(10.18 \text{ million Km}^2) = 16.136$, and the minimum value represents $\log(18\,200 \text{ Km}^2) = 9.81$.

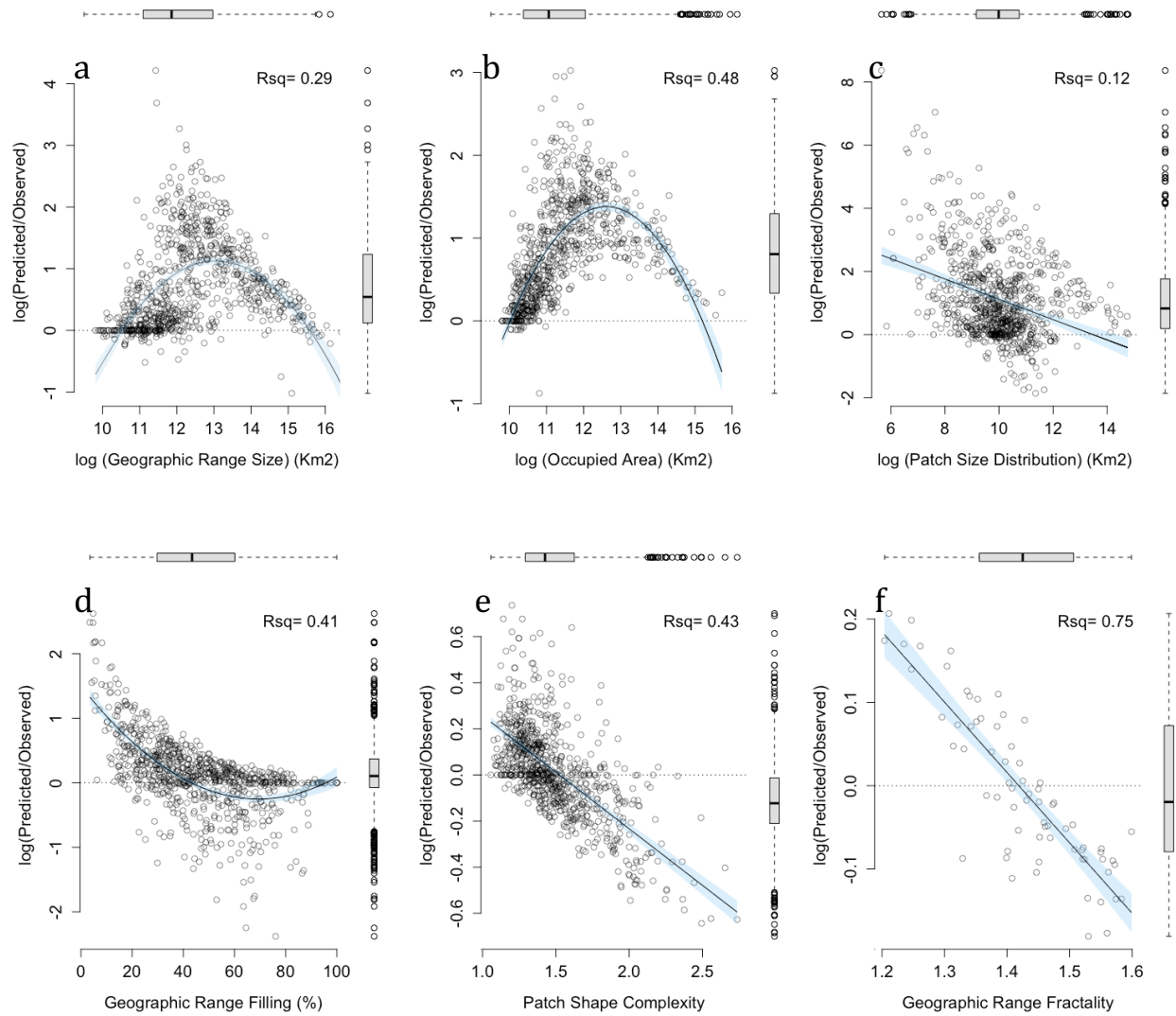


Figure 4. The relationship between the geographic range size and structure metrics quantified using observed occurrence data (x axis) and the log response ratio between the metrics calculated from habitat suitability maps and from observed distributions (y axis). The zero line indicates no difference between the observed and predicted metric. R^2 = R squared value of the linear regression model. The scale of the y axes differs between different metrics. Horizontal and vertical boxplots represent summary values on the x and y axes respectively.

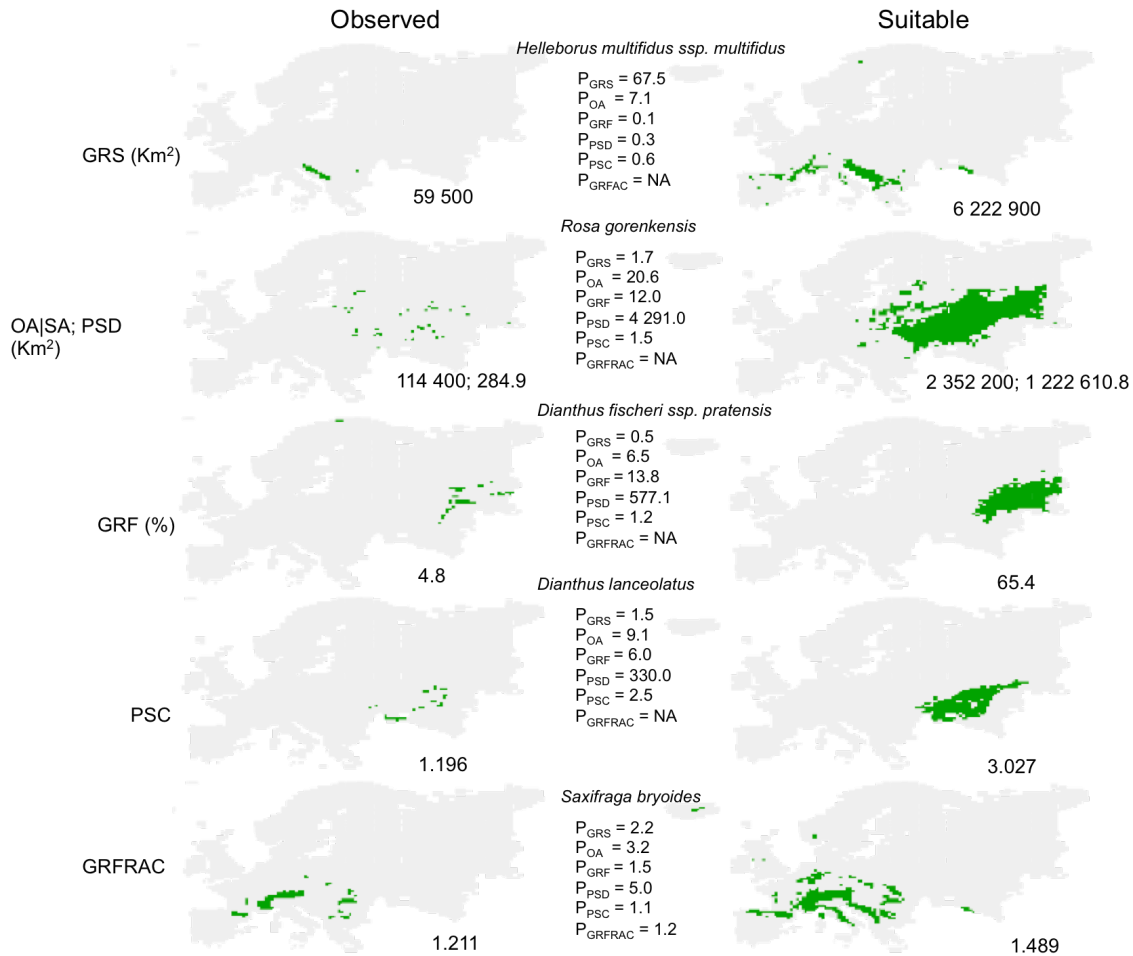


Figure 5. Examples of geographic range size and structure metrics quantified from habitat patches occupied and predicted suitable, and the proportion (P) between the geographic range structure metrics quantified from habitat patches predicted suitable and occupied patches in five example species. Numbers next to the maps indicate the value of the geographic range structure metric. Species were selected to illustrate the largest proportion between predicted and occupied metric (P) from a pool of 813 species endemic to Europe. Proportions of different range structure metrics differ within the same species and between species.

Values above 1 indicate larger values of the predicted relative to the observed pattern, values below 1 indicate larger values of the observed relative to the predicted pattern. GRS = geographic range size, OA|SA = occupied (suitable) area, GRF = geographic range filling, PSD = patch size distribution, PSC = patch shape complexity, GRFRAC = geographic range fractality.

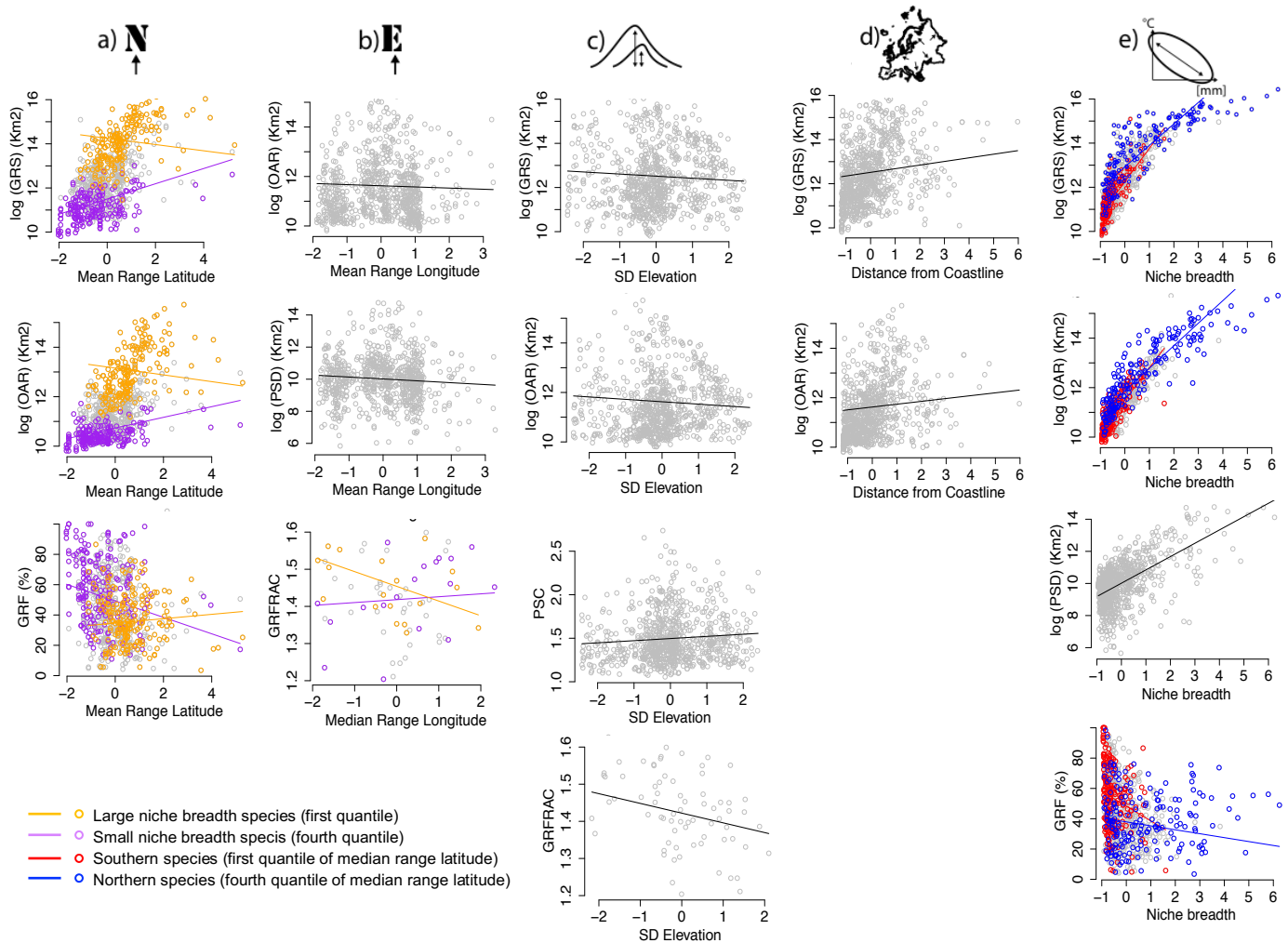


Figure 6. Scatterplots showing the significant effects of **a)** median range latitude, **b)** median range longitude, **c)** topographic heterogeneity (SD elevation), **d)** distance from coastline and **e)** niche breadth on the geographic range structure metrics calculated from observed distributions, modeled using PGLS. GRS=geographic range size, OA=occupied area, PSD=patch size distribution, GRF=geographic range filling, PSC=patch shape complexity, GRFRAC=geographic range fractality. Points represent species, and fitted lines represent the conditional effect of the individual variable, holding other variables constant. Values on the x axes are centered on zero with unit variance. Effect sizes and model details are presented in Supporting material.

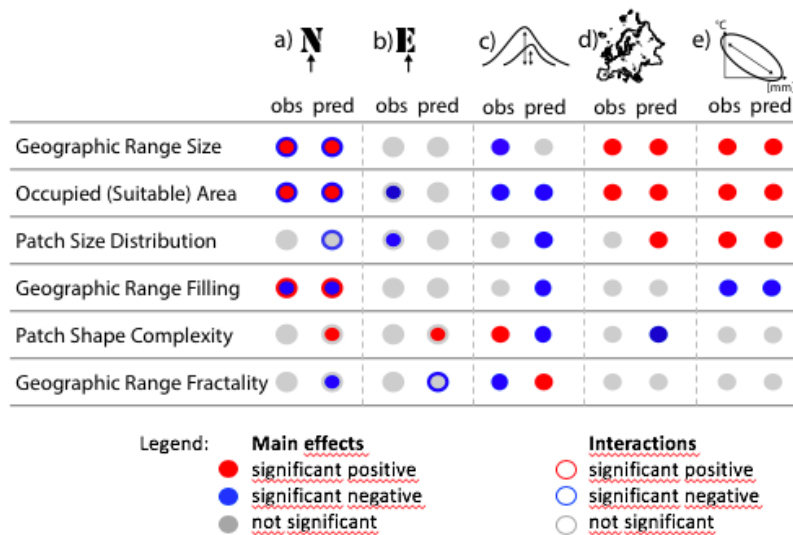


Figure 7. Diagram showing significant effects of **a)** Median range latitude, **b)** Median range longitude, **c)** Topographic heterogeneity (SD Elevation) **d)** Distance from coastline, and **e)** Niche breadth on six geographic range structure metrics modeled using PGLS on species' observed and predicted distributions. Effect sizes and model details are presented in Figure 6 and Supporting material.