

Model linearity breeds contempt: using non-linear models to uncover general biogeographical patterns

Bernat Bramon Mora^{1,*} and Jake M. Alexander¹

¹Institute of Integrative Biology, ETH Zürich, Zürich, Switzerland; *bernat.bramon@gmail.com

1 Abstract

2 We know a lot about the factors that could theoretically influence species' distributions,
3 and a rapidly growing body of research have been primarily focused on trying to untangle
4 some of such biotic and abiotic predictors—with an increasing effort placed in improving
5 the predictive power of statistical models. However, much less is known about how species'
6 distributions compare to each other. Here, we use a conceptually more conservative ap-
7 proach to instead understand and compare basic aspects regarding the shape of species'
8 distribution along environmental gradients.

9 Introduction

10 One of the central goals of ecology is to understand the ways species are distributed across
11 space and time (ref). Over the last two decades, ecologists have developed multiple distri-
12 bution models to try to untangle the factors that play a role in defining such distributions
13 (Guisan & Zimmermann, 2000). These models estimate species' realized niches using sev-
14 eral covariates, including environmental variables (?), species ecological traits' (Pollock
15 *et al.*, 2012) and phylogenetic relations (Ives & Helmus, 2011). More recently, some of the
16 focus have shifted towards approaches that estimate and account for biotic factors, such
17 as competitive or facilitative relationships between species (Ovaskainen *et al.*, 2017). The
18 idea is that by untangling the ways in which such biotic and abiotic factors shape species'
19 distributions, we can gain a mechanistic understanding on how ecological communities are

20 established and change over time. However, while these factors can increase the predictive
21 performance of some of the models (Norberg *et al.*, 2019), the interpretation of the corre-
22 sponding parameter estimates has been recently questioned (Harris, 2016; Thurman *et al.*,
23 2019; ?). This was best illustrated by Blanchet *et al.* (2020), who used basic statistical ar-
24 guments to highlight the artefactual nature of the link between co-occurrence and species'
25 ecological interactions drawn by some distribution models.

26 The value of gaining a mechanistic understanding of species' distributions is unques-
27 tionable (ref), with several studies highlighting the importance of factors such as biotic
28 interactions and dispersal ability in setting species' range limits (Wisz *et al.*, 2013; Pol-
29 lock *et al.*, 2014; Neuschulz *et al.*, 2018). That said, a lot can be learned from taking
30 a phenomenological approach, focussing instead on the description of basic properties of
31 species' realized niches. For example, the study of species' range sizes along environmental
32 gradients can reveal general biodiversity patterns that are crucial from a conservation and
33 management perspective (Stevens, 1992). Differences in species' responses to the environ-
34 ment could shed light on how climatic processes and historical contingencies have shaped
35 their distributions (Rohde, 1992; ?). Other properties, such as the skewness of species'
36 distributions, can also reveal general underlying processes regarding species' physiological
37 tolerance to different environmental conditions (Kaufman, 1995). More generally, under-
38 standing the shape of species' realized niches and the extend to which these vary across
39 species is a crucial issue in ecology and biogeography (ref); however, we do not have an
40 effective way to parsimoniously compare the realized niches of many species. Indeed, there
41 is no general agreement on the shape of species' distributions (ref).

42 Many ecological textbooks (Krebs, 1972) assume the shape of species distributions to
43 be unimodal and symmetric, but some have warned that empirical distributions can take
44 many different forms (Austin, 1987; ?). In practice, distribution frameworks often use lo-
45 gistic regressions with a linear relationship between covariates (but see XX and YY). This
46 is useful because it simplifies the optimization process, but it comes with some statistical
47 shortcomings. First and foremost, such response curve and the linear relationship between
48 covariates often comes with a set of implicit mathematical constraints that might not be bio-
49 logically justified. From a purely statistical perspective, if all that we are willing to assume

is that species occupy finite geographic ranges—i.e. their probability distributions have finite variance—the most conservative statistical approach is to model these as a Gaussian distributions (Frank, 2009). This is rarely the starting point in most statistical frameworks that study general biodiversity patterns (but see ref), choosing to use instead Gaussian-logit response curves (refs). Other factors might then condition species distributions to showcase fat-tails or a skewed shapes, revealing interesting ecological processes shaping biodiversity patterns (Austin, 1976; Minchin, 1987). The starting point, nevertheless, should be the one that makes the fewest assumptions (i.e. the maximum entropy distribution; Frank 2009), and every new shape will imply a hypotheses on how communities are distributed (D’Amen *et al.*, 2017). Second, the aforementioned structural constraints also limit our ability to include any prior information to our parameter estimates. Observations on species’ geographic variation and optimal climatic conditions have long been documented, with extensive databases compiled by botanists and field ecologists documenting basic knowledge on species’ realized niches (e.g. Landolt *et al.* 2010). That said, this information is rarely accounted for in most modelling approaches, potentially because there is not a straightforward way to feed this information into the parameters of a linear model (Scherrer & Guisan 2019; but see ter Braak & Looman 1986). Finally, and perhaps most importantly, a direct biological interpretation of parameter estimates in linear models becomes increasingly difficult as one moves from unimodal and symmetric distributions (ter Braak & Looman, 1986; Jamil & ter Braak, 2013) to skewed distributions (Huisman *et al.*, 1993), making the tests of hypothesis on global biodiversity patterns particularly challenging. For example, Huisman *et al.* (1993) proposed several non-linear models to characterize several features of species’ response curves; however, species’ environmental indicator values, range size or distribution skewness are difficult to understand altogether following these model structures.

The field of ecology has quickly moved towards mechanistic and process-based approaches to understand species’ distributions (Warton *et al.*, 2015). This has resulted in a plethora of models accounting for several biotic and abiotic factors into the predictions of species co-occurrence. Here, we instead rethink traditional modelling approaches and develop a conceptually simple—and yet statistical and computationally complex—statistical framework to revisit some classic hypothesis in ecology and biogeography. In particular, we develop a Bayesian hierarchical model that accounts for all prior information that we have regarding

81 the distribution of alpine plant species along an elevation gradient in the Swiss Alps, in-
 82 cluding expert knowledge on species environmental indicator values, range sizes, and plant
 83 physiology. We start by considering species' response curves as Gaussian distributed, and
 84 then we adapt our model to allow for skewed and long-tailed distributions. Using this
 85 statistical framework, we are able to compare the basic properties of the realized niches of
 86 multiple species, testing for the existence of general biogeographical patterns. First, we test
 87 for the Rapoport's rule, which predicts a positive relationship between range size and eleva-
 88 tion (Stevens, 1992). While this pattern has been largely studied for multiple systems and
 89 across gradients (McCain & Knight, 2013); contrasting evidence suggests this rule not to
 90 be pervasive across species (Ribas & Schoereder, 2006; Bhattarai & Vetaas, 2006; McCain
 91 & Knight, 2013). Our results not only allow us to properly test the existence of this geo-
 92 graphical pattern, but they also showcase variation in how different types of species, such
 93 as native or neophytes, might respond to an environmental gradient. Second, we study
 94 whether or not species' distributions show steeper declines towards stressful conditions,
 95 testing the so-called abiotic stress limitation hypothesis (ref). Normand *et al.* (2009) tested
 96 this for vegetation data using Huisman *et al.*'s statistical models for several independent
 97 species, finding no clear support for such a hypothesis. Our results are able to shed light on
 98 this geographical pattern as well as to highlight the degree to which different species will
 99 showcase different levels of decline towards stressful conditions. Specifically, we are able to
 100 link plant physiological traits to the skewness of their distributions. Overall, we use models
 101 that are solely constrained by the empirical information that we truly have regarding our
 102 system, relaxing as much as possible the structural constraints of the statistical framework.
 103 Using these models, we are able to uncover the approximate shape of empirical plant distri-
 104 butions and answer fundamental questions regarding the way systems of many species are
 105 distributed along environmental gradients.

Methods

Empirical data

We studied the distribution of alpine plant communities along an elevation gradient. To do so, we combined two different datasets: i) one describing the co-occurrence of species across multiple open grasslands in the Swiss Alps, and ii) an extensive floristic database containing environmental and physiological traits for all vegetation across Switzerland (Landolt *et al.*, 2010).

Distribution data

We used data describing the distribution of 798 species across 912 sites covering most of the mountain region of the Western Alps in the Canton de Vaud (Switzerland; Scherrer & Guisan 2019). Each of these sites is a 8×8 m plot placed somewhere along an elevation range from 375 m to 3210 m. In all sites, presence/absence data as well as Braun-Blanquet abundance-dominance classes were recorded for all species. Additionally, we used meteorological data provided by Scherrer & Guisan (2019), containing multiple variables characterizing the climate in each site at high spatial resolution (25 m). This dataset was compiled based on 30 years (1961–1990) of records from national weather stations. Since most of the data is highly correlated, we calculated the main axes of variation of the following scaled variables: daily minimum, maximum and average temperature; sum of growing degree-days above 5°C ; mean temperature of wettest quarter; annual precipitation, precipitation seasonality, and precipitation of driest quarter (see Supplementary Methods; Supplementary Fig. 1).

Floristic data

To complement the aforementioned distribution data, we used a floristic database of most vegetation across Switzerland. This database was build based on expert knowledge and field experience of botanists and ecologists, and contains information regarding species' environmental preferences and physiological traits. Species' environmental preferences in

132 this database can be used to inform distribution models—e.g. as an informative prior in
133 a Bayesian framework. These are characterized following the ecological indicator values
134 developed by Landolt *et al.* (2010), providing both an estimate of the average conditions
135 in which a species can be found as well as a broad description of their range of variation.
136 These values are provided for a range of 10 climatic variables, including temperature, con-
137 tinentality, light conditions, as well as moisture, acidity and nutrient content of the soil (see
138 a full list and description of the ecological indicators in the Supplementary Table 1; Landolt
139 *et al.* 2010). On the other hand, the information regarding species’ physiological traits rep-
140 resents general descriptions of species’ growth and life strategies—examples include their
141 growth forms, nature of the storage organs, dispersal ability and pollinator agents. In total,
142 we identify more than 120 binary traits that characterize the physiology of species (see a
143 full list and description of the ecological indicators in the Supplementary Table 1; Landolt
144 *et al.* 2010). Finally, and in addition to species’ environmental preferences and physiological
145 traits, the floristic data also contains information on species types (e.g. identifying those
146 species that are neophytes)—and growth tendency (e.g. indicating species that have shown
147 decline or increase in their populations over the recent years). We describe this information
148 in more detail in the Supplementary Table 1.

149 **Baseline model**

150 There is a long list of model structures well suited to characterize species’ distributions (see
151 Norberg *et al.* 2019). As a baseline model, however, we were interested in a hierarchical
152 model that does not make any assumptions regarding the shape of the distributions, and
153 yet explicitly incorporates all information that we have regarding plant’s environmental
154 preferences. More specifically, we wanted to account for the climatic indicator values and
155 range of variation registered in the floristic database for all plants in our dataset. These
156 two values provide basic information regarding plant’s optimal environmental conditions
157 and width of their distributions.

158 *Response curve*

159 To choose an appropriate response curve, we first need to agree on what we truly know
160 about the system. Given the prior information that we have about the system, we know
161 that species occupy specific geographic ranges; therefore, we know that their distributions
162 have finite variance. While we could also assume that many other factors might influence
163 species' presence in a given site—e.g. the biotic interactions among species in the site—we
164 do not necessarily have an *a priori* expectation of how exactly these factors will influence
165 the shape of species' distributions. Therefore, for this baseline model, if all that we are
166 willing to assume about species' realized niches is that these have finite variance, the most
167 conservative assumption and the safest bet—i.e. the one with the largest entropy—is that
168 they follow a Gaussian distribution. That is, given the presence/absence or abundance y_{ij}
169 of any species i in any given site j , and an environmental variable x_j , we define species'
170 responses to the environment as

$$y_{ij} \sim F(p_{ij})$$
$$\log(p_{ij}) = -\alpha_i - \gamma_i(x_j - \beta_i)^2, \quad (1)$$

171 where F is the likelihood function, and α_i , β_i^k , and γ_i describe amplitude of the probability
172 p_{ij} , species' average climatic suitability and range of variation along the an environmental
173 gradient, respectively. Notice that F characterizes a Bernoulli distribution when considering
174 binary data, and it characterizes an ordered categorical likelihood function when we consider
175 Braun-Blanquet abundance-dominance classes as response variables (see the full description
176 of both models in the Supplementary Methods). For the sake of simplicity, we use only
177 one environmental variable to characterize species' probability distribution. That said,
178 this model can easily be generalized to account for multiple predictors (see Supplementary
179 Methods).

180 *Model priors*

181 The model structure described above allows us to explicitly incorporate all prior knowledge
182 that we have regarding species' distributions contained in the floristic database. To do so,

183 we define the prior distributions for the parameters in model (1) as:

$$\begin{aligned}
\beta_i &\sim \text{MVNormal}(\hat{\beta}, \Sigma^\beta) \\
\log(\gamma_i) &\sim \text{MVNormal}(\hat{\gamma}, \Sigma^\gamma) \\
\log(\alpha_i) &\sim \text{Normal}(\hat{\alpha}, \sigma_\alpha) \\
\hat{\beta}, \hat{\gamma}, \hat{\alpha} &\sim \text{Normal}(0, 1) \\
\sigma_\alpha &\sim \text{Exponential}(1)
\end{aligned} \tag{2}$$

184 where parameters γ_i and β_i are expressed as multivariate normal distributions—i.e. Gaus-
185 sian processes—such that Σ^β and Σ^γ are variance-covariance matrices describing species’
186 similarity in terms of their average climatic suitability and range of variation along the dif-
187 ferent environmental gradients, respectively. We define these variance-covariance matrices
188 as follows:

$$\Sigma_{ij} = \eta \exp(-\rho D_{ij}^2) + \delta_{ij} \sigma, \tag{3}$$

189 where Σ_{ij} characterizes the covariance between any pair of species i and j , and δ_{ij} is the
190 Kronecker delta. Notice that such a covariance structure declines exponentially with the
191 square of a distance matrix D_{ij} , which characterize differences between species computed
192 using our prior information. In the floristic database, this information is represented by the
193 set of ordinal specified for the different species. While there are many different ways to turn
194 ordinal data into distance matrices, we choose to use a mixed-membership stochastic block
195 model because it allows us to deal with cases of missing data (see Supplementary Methods
196 for extended details; Godoy-Lorite *et al.* 2016). In each covariance matrix, the hyperparam-
197 eter ρ determines the rate of decline of the covariance between any two species, and η defines
198 its maximum value. The hyperparameter σ describes the additional covariance between the
199 different observations for any given species. For all these hyperparameters, we choose weekly
200 informative priors such that $\sigma, \eta \sim \text{Exponential}(1)$ and $\rho \sim \text{Exponential}(0.5)$.

201 *Alternative variance-covariance structures*

202 The model structure defined above allows us to test how different sources of information
203 characterize each of the different parameters. Specifically, we can do this by modifying

Eq. (3). For example, imagine that we have multiple matrices D^k characterizing species' differences along different axis of variation—e.g. two matrices characterizing physiological and environmental traits. One can modify Eq. (3) for a particular parameter—e.g. β_i —such that

$$\Sigma_{ij} = \eta \exp \left(- \sum_k \rho_k D_{ij}^k{}^2 \right) + \delta_{ij} \sigma, \quad (4)$$

where now ρ_k are separate relevance hyperparameters for each distance matrix in the total variance of β_i .

Sampling the posterior

We generated the posterior samples for the Bayesian models with the Hamiltonian Monte Carlo algorithm implementation provided by the R package ‘rstan’ to (Stan Development Team, 2021). Sampling models like the ones described above can be computationally very expensive. This is especially true when using ordered categorical likelihood functions (see Stan Development Team 2021). Therefore, we focus on those species for which we have more than 10 occurrences when modelling binary data, which is the case for the majority of the results of this work. When using ordinal data, we limit our study to those species for which have more than 50 occurrences.

To test the performance of the model as well as our choice of prior distributions, we modelled simulated data and compared the sampled posterior distributions to the data-generating parameters (see Supplementary Methods; Supplementary Fig. 2). Notice that using the link function in Eq. (1) could cause problems when sampling the model, and some adjustments need to be made when specifying the model (see Supplementary Methods and the Code Availability section).

Modifying the baseline model

We proposed a baseline model that is naive regarding how the data is distributed, and yet accounts for all prior information that we have about the system. Now, we want to modify this model to test the extent to which empirical species' distributions showcase different properties, while preserving both the interpretation of the parameter estimates and the

230 prior information. More specifically, to propose new species' response curves, we follow
 231 three key criteria: (i) the probability distribution must have a defined variance and mean,
 232 (ii) the Gaussian shape must be a special case of the probability distribution, and (iii) there
 233 must be a re-parametrization of the model that allows us to keep the same prior information
 234 and interpretable parameters.

235 *Fat-tail response curve*

236 Distributions with fat-tails are very common across fields, as they can capture processes such
 237 as seasonality (e.g. in communications patterns; [Malmgren et al. 2008](#)) or some stochastic
 238 events (e.g. in the spread of infectious diseases; [Wong & Collins 2020](#)). Indeed, fat-tail
 239 distributions are pervasive in ecology; for example, species' dispersal patterns have been
 240 shown to have fat-tails due to natural variability among individuals ([Petrovskii et al., 2009](#)).
 241 Therefore, one might expect these properties to also emerge in alpine communities, where
 242 seasonality and dispersal patterns are crucial factors determining species' distributions (?).
 243 To accommodate this feature into our baseline model, we could consider response curves
 244 that follow a generalized error or a non-standardized Student's t-distributions. In both
 245 cases, the normal distribution is a special case of them, and both contain parameters that
 246 regulate the kurtosis of the probability distribution (Supplementary Figure XX).

247 Starting with the generalized error distribution, we can adapt Eq. (1) to present this
 248 non-linear form as follows:

$$\log(p_{ij}) = -\alpha_i - \gamma'_i |x_j - \beta_i|^{\nu_i}, \quad (5)$$

249 where $\gamma'_i = g(\gamma_i, \nu_i)$, and ν_i is a parameter that describes the kurtosis of the distribution,
 250 which we define as $\nu_i \in (1, \infty)$. Following this, we choose an adaptive prior for this set
 251 of new parameter such that $\log(\nu_i - 1) \sim \text{Normal}(\hat{\nu}, \sigma_\nu)$, where $\hat{\nu} \sim \text{Normal}(0, 1)$ and
 252 $\sigma_\nu \sim \text{Exponential}(2)$. Given the relationship between γ'_i and γ_i , we can re-parametrize the
 253 model and follow Eq. (2) to define the corresponding prior distributions (see Supplementary
 254 Table 2; Code Availability section; ?). Notice that the Gaussian distribution will naturally
 255 emerge when $\nu_i = 2$.

256 In a similar fashion, we can use a non-standardized Student's t-distribution by modifying
 257 Eq. (1). In this case, the response curve can be characterized as follows:

$$p_{ij} = \alpha'_i \left(1 + \frac{\gamma'_i}{\nu_i} (x_j - \beta_i)^2 \right)^{-\frac{\nu_i+1}{2}}, \quad (6)$$

258 where $\alpha'_i = f_1(\alpha_i, \gamma'_i, \nu_i)$, $\gamma'_i = f_2(\gamma_i, \nu_i)$, and ν_i is a parameter that describes the kurtosis of
 259 the distribution, which we define as $\nu_i \in (2, \infty)$. These boundary conditions are necessary
 260 in order for the distributions to have finite variance. Similar to the description above, we
 261 choose an adaptive prior such that $\log(\nu_i - 2) \sim \text{Normal}(\hat{\nu}, \sigma_\nu)$, where $\hat{\nu} \sim \text{Normal}(0, 1)$
 262 and $\sigma_\nu \sim \text{Exponential}(1)$. Again, we can re-parametrize the model following f_1 and f_2 ,
 263 and set the rest of the prior distributions as above (see Supplementary Table 2; Code
 264 Availability section). In this case, the Gaussian distribution is a special case of Eq. (6)
 265 when $\nu_i \rightarrow \infty$ (?).

266 *Skewed response curve*

267 When species experience abiotic or biotic pressures that increase or decrease along an en-
 268 vironmental gradient, one might expect their distributions to be skewed in one direction.
 269 Likewise, this same skewed shape can also be the result of asymmetric environmental tol-
 270 erance. One way to accommodate this feature to our models is by considering a skewed
 271 normal distribution (Supplementary Figure XX). As for the two cases described above, the
 272 Gaussian is a special case of this distribution, and it contains a parameter that controls
 273 for the level and direction of 'skewness'. Importantly, this distribution presents normal-like
 274 tails; therefore, the added skewness does not make additional assumptions regarding how
 275 species' are distribution along the gradient. To test for the existence of this feature, we
 276 modified the species' response curve in Eq. (1) as

$$p_{ij} = \hat{p}_{ij} \left[1 + \text{erf} \left(\lambda_i (x_j - \beta'_i) \sqrt{\frac{\gamma'_i}{2}} \right) \right]$$

$$\log(\hat{p}_{ij}) = -\alpha'_i - \gamma'_i (x_j - \beta'_i)^2, \quad (7)$$

277 where $\alpha'_i = q_1(\alpha_i, \gamma'_i, \lambda_i)$, $\beta'_i = q_2(\gamma'_i, \lambda_i)$, $\gamma'_i = q_3(\gamma_i, \lambda_i)$, and λ_i is a parameter that
 278 describes the skewness of the distribution, which we define as $\lambda_i \in (-10, 10)$. Notice that

these boundary conditions for λ_i are arbitrarily set for computational purposes, as this parameter is theoretically not bounded. The function ‘erf’ is the error function (see ?). We choose an adaptive prior such that $\lambda_i \sim \text{Normal}(\hat{\lambda}, \sigma_\lambda)$, where $\hat{\lambda} \sim \text{Normal}(0, 1)$ and $\sigma_\lambda \sim \text{Exponential}(1)$. This model can be re-parametrized following q_1 , q_2 and q_3 , and set the rest of the prior distributions as above (see Supplementary Table 2; Code Availability section). In this case, the Gaussian distribution is a special case of Eq. (7) when $\lambda_i = 0$ (?).

One way to test the extend to which species’ physiological traits inform the skewness of the distribution of any species i is by characterizing λ_i as a Gaussian process. That is, we can consider the prior distribution for λ_i as a multivariate normal with a variance covariance matrix Σ^λ . As described before, this variance-covariance matrix can be estimated by the model following Eq. (3), and its structure can shed light on how informative the physiological traits are. Likewise, we could instead use Eq. (4) and test the effectiveness of other prior information in explaining the skewness of species’ distributions.

Results

We studied the distribution data to characterize species’ realized niches. To do so, we selected the main axis of variation of all environmental variables (Supplementary Fig. XX). Then, using the presence or absence of species per site as response variable, we sampled the posterior distributions of the baseline model, accounting for the information in the floristic database regarding species’ indicator values and range of variation. Studying the relationship between the mean and variance of the distributions, we found clear evidence of the Rapapor’s rule. Specifically, we found that species’ range sizes decline with elevation and temperature (i.e. β_i and γ_i in the baseline model are negatively correlated; Fig. 1). The relationship was also found when using instead ordinal data (Supplementary Fig. XX), and it does not seem to depend on species’ type or reflect species’ abundance change tendency over the years (Supplementary Fig. XX). This pattern was not present along the second axis of variation of the environmental variables (Supplementary Fig. XX). The comparison between the other parameter estimates revealed additional, somewhat more expected, relationships. In particular, we found the amplitude of distributions to be positively and

negatively correlated with the mean and the variance along the elevation gradient, respectively (Supplementary Fig. XX).

Maintaining the symmetry of species' distributions, we then allowed the kurtosis—or shape of the tails—of these to vary in different ways. To do so, we changed the response curve of our Bayesian model to follow either a generalized error distribution (Eq. 5) or a non-standardized Student's t-distribution (Eq. 6).

We then studied the level of information contained in physiological traits to explain the properties of distributions. Specifically, using Eq. (4), we accounted for both physiological and environmental traits in the variance-covariance structure for the mean and standard deviation of distributions. As expected, the posterior distributions also showcase a positive relationship between... This is expected as ... Likewise, studying ... Following this, we added additional ... using Eq. (4). When studying plant...

Discussion

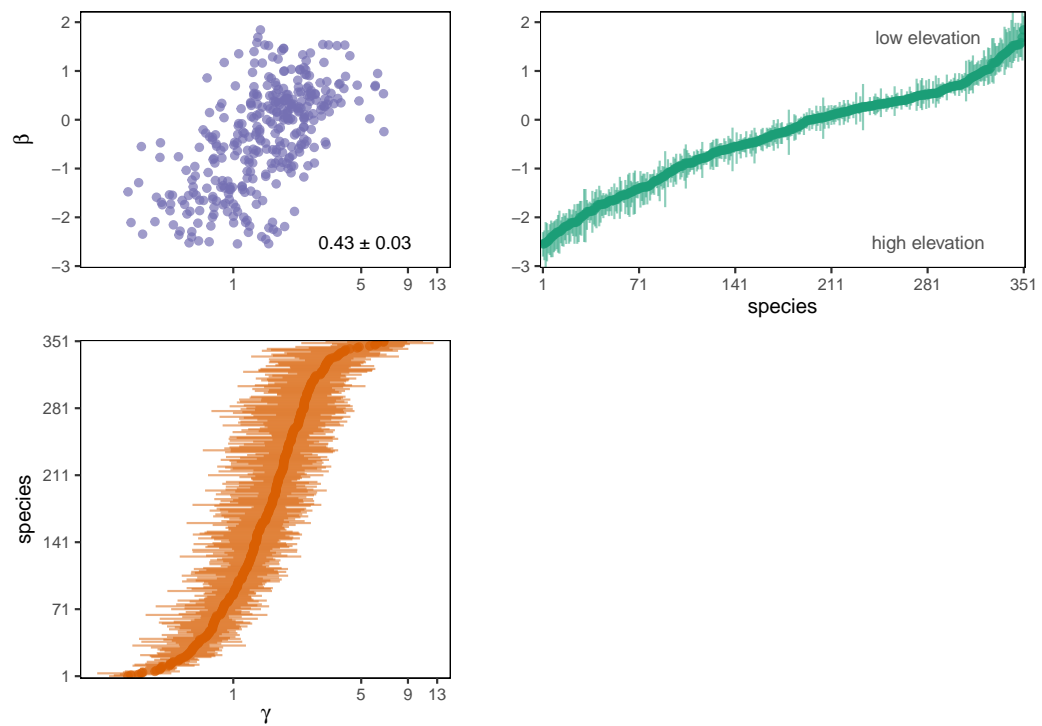


Figure 1: Relationship between mean and variance of species' distributions. These are the results for the main axis of variation for the climatic data (results for the second axis of variation presented in the Supplementary Fig. 2).

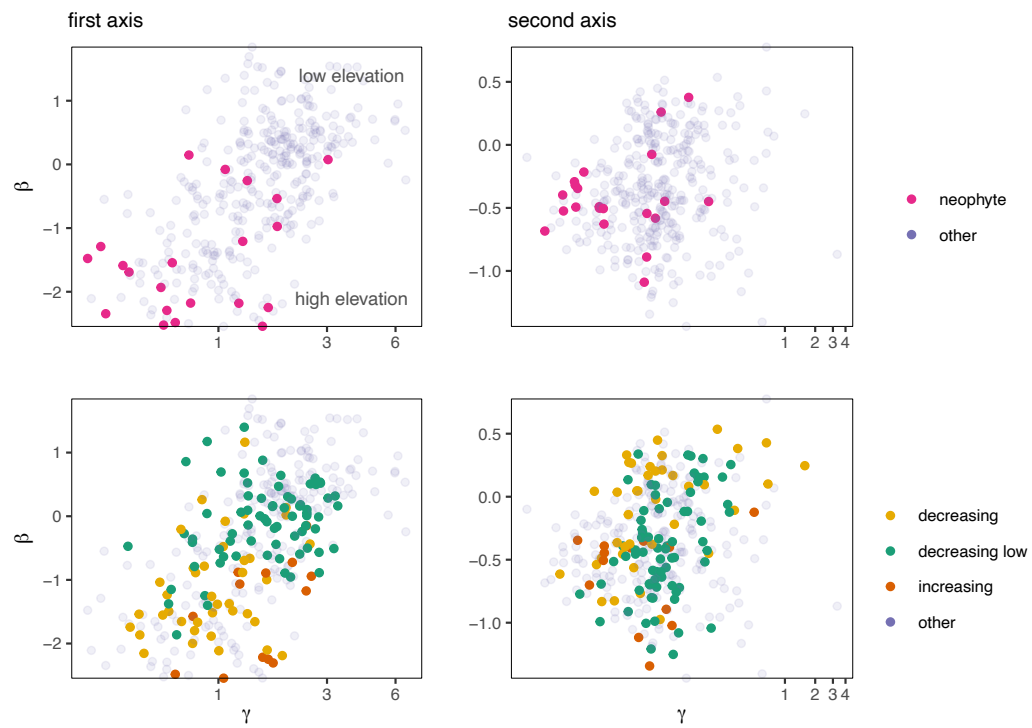


Figure 2: Are there clear geographical patterns for neophytes and for species with decreasing or increasing abundance?

References

- Austin, M. (1976). On non-linear species response models in ordination. *Vegetatio*, 33, 33–41.
- Austin, M. P. (1987). Models for the analysis of species' response to environmental gradients. *Vegetatio*, 69, 35–45.
- Bhattacharai, K. R. & Vetaas, O. R. (2006). Can Rapoport's rule explain tree species richness along the Himalayan elevation gradient, Nepal? *Diversity and Distributions*, 12, 373–378.
- Blanchet, F. G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, 23, 1050–1063.
- D'Amen, M., Rahbek, C., Zimmermann, N. E. & Guisan, A. (2017). Spatial predictions at the community level: From current approaches to future frameworks. *Biological Reviews*, 92, 169–187.
- Frank, S. A. (2009). The Common Patterns of Nature. *Journal of evolutionary biology*, 22, 1563–1585.
- Godoy-Lorite, A., Guimerà, R., Moore, C. & Sales-Pardo, M. (2016). Accurate and scalable social recommendation using mixed-membership stochastic block models. *Proceedings of the National Academy of Sciences*, 113, 14207–14212.
- Gotelli, N. J. & Ulrich, W. (2010). The empirical Bayes approach as a tool to identify non-random species associations. *Oecologia*, 162, 463–477.
- Guisan, A. & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–186.
- Harris, D. J. (2016). Inferring species interactions from co-occurrence data with Markov networks. *Ecology*, 97, 3308–3314.
- Huisman, J., Olff, H. & Fresco, L. F. M. (1993). A hierarchical set of models for species response analysis. *Journal of Vegetation Science*, 4, 37–46.

- Ives, A. R. & Helmus, M. R. (2011). Generalized linear mixed models for phylogenetic analyses of community structure. *Ecological Monographs*, 81, 511–525.
- Jamil, T. & ter Braak, C. J. F. (2013). Generalized linear mixed models can detect unimodal species-environment relationships. *PeerJ*, 1, e95.
- Kaufman, D. M. (1995). Diversity of New World Mammals: Universality of the Latitudinal Gradients of Species and Bauplans. *Journal of Mammalogy*, 76, 322–334.
- Krebs, C. J. (1972). *Ecology: The Experimental Analysis of Distribution and Abundance/by Charles J. Krebs*. 4th edn. Harper & Row, New York.
- Landolt, E., Bäumler, B., Ehrhardt, A., Hegg, O., Klötzli, F., Lämmli, W., Nobis, M., Rudmann-Maurer, K., Schweingruber, F. H., Theurillat, J.-P., Urmi, E., Vust, M. & Wohlgemuth, T. (2010). *Flora indicativa: Ökologische Zeigerwerte und biologische Kennzeichen zur Flora der Schweiz und der Alpen*. Haupt, Bern. ISBN 978-3-258-07461-0.
- Malmgren, R. D., Stouffer, D. B., Motter, A. E. & Amaral, L. A. N. (2008). A Poissonian explanation for heavy tails in e-mail communication. *Proceedings of the National Academy of Sciences*, 105, 18153–18158.
- McCain, C. M. & Knight, K. B. (2013). Elevational Rapoport’s rule is not pervasive on mountains. *Global Ecology and Biogeography*, 22, 750–759.
- Minchin, P. R. (1987). An evaluation of the relative robustness of techniques for ecological ordination. *Vegetatio*, 69, 89–107.
- Neuschulz, E. L., Merges, D., Bollmann, K., Gugerli, F. & Böhning-Gaese, K. (2018). Biotic interactions and seed deposition rather than abiotic factors determine recruitment at elevational range limits of an alpine tree. *Journal of Ecology*, 106, 948–959.
- Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., Araújo, M. B., Dallas, T., Dunson, D., Elith, J., Foster, S. D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O’Hara, B., Hill, N. A., Holt, R. D., Hui, F. K. C., Husby, M., Kålås, J. A., Lehtikainen, A., Luoto, M., Mod, H. K., Newell, G., Renner, I., Roslin, T., Soininen, J., Thuiller, W., Vanhatalo, J., Warton, D., White, M., Zimmermann, N. E., Gravel, D.

374 & Ovaskainen, O. (2019). A comprehensive evaluation of predictive performance of 33
375 species distribution models at species and community levels. *Ecological Monographs*, 89,
376 e01370.

377 Normand, S., Treier, U. A., Randin, C., Vittoz, P., Guisan, A. & Svenning, J.-C. (2009).
378 Importance of abiotic stress as a range-limit determinant for European plants: Insights
379 from species responses to climatic gradients. *Global Ecology and Biogeography*, 18, 437–
380 449.

381 Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F. G., Duan, L., Dunson, D., Roslin,
382 T. & Abrego, N. (2017). How to make more out of community data? A conceptual
383 framework and its implementation as models and software. *Ecology Letters*, 20, 561–576.

384 Petrovskii, S., Morozov, A., Taylor, A. E. P. D. & DeAngelis, E. D. L. (2009). Dispersal
385 in a Statistically Structured Population: Fat Tails Revisited. *The American Naturalist*,
386 173, 278–289.

387 Pollock, L. J., Morris, W. K. & Vesk, P. A. (2012). The role of functional traits in species
388 distributions revealed through a hierarchical model. *Ecography*, 35, 716–725.

389 Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O’Hara, R. B., Parris, K. M., Vesk,
390 P. A. & McCarthy, M. A. (2014). Understanding co-occurrence by modelling species
391 simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology*
392 *and Evolution*, 5, 397–406.

393 Ribas, C. R. & Schoereder, J. H. (2006). Is the Rapoport effect widespread? Null models
394 revisited. *Global Ecology and Biogeography*, 15, 614–624.

395 Rohde, K. (1992). Latitudinal Gradients in Species Diversity: The Search for the Primary
396 Cause. *Oikos*, 65, 514–527.

397 Scherrer, D. & Guisan, A. (2019). Ecological indicator values reveal missing predictors of
398 species distributions. *Scientific Reports*, 9, 1–8.

399 Stan Development Team (2021). RStan: The R interface to Stan.

400 Stan Development Team (2021). Stan Modeling Language Users Guide and Reference
401 Manual.

402 Stevens, G. C. (1992). The Elevational Gradient in Altitudinal Range: An Extension of
403 Rapoport’s Latitudinal Rule to Altitude. *The American Naturalist*, 140, 893–911.

404 ter Braak, C. J. F. & Looman, C. W. N. (1986). Weighted averaging, logistic regression
405 and the Gaussian response model. *Vegetatio*, 65, 3–11.

406 Thurman, L. L., Barner, A. K., Garcia, T. S. & Chestnut, T. (2019). Testing the link
407 between species interactions and species co-occurrence in a trophic network. *Ecography*,
408 42, 1658–1670.

409 Warton, D. I., Blanchet, F. G., O’Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C.
410 & Hui, F. K. C. (2015). So Many Variables: Joint Modeling in Community Ecology.
411 *Trends in Ecology & Evolution*, 30, 766–779.

412 Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dor-
413 mann, C. F., Forchhammer, M. C., Grytnes, J.-A., Guisan, A., Heikkinen, R. K., Høye,
414 T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.-C., Normand, S., Öckinger, E.,
415 Schmidt, N. M., Termansen, M., Timmermann, A., Wardle, D. A., Aastrup, P. & Sven-
416 ning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised
417 assemblages of species: Implications for species distribution modelling. *Biological Re-*
418 *views*, 88, 15–30.

419 Wong, F. & Collins, J. J. (2020). Evidence that coronavirus superspreading is fat-tailed.
420 *Proceedings of the National Academy of Sciences*, 117, 29416–29418.