

Model simplicity breeds contempt: using simple models to answer basic questions on species' distributions

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 (??). These models estimate species' realized niches using several covariates, including en-
14 vironmental variables (?), species ecological traits' (Pollock *et al.*, 2012) and phylogenetic
15 relations (Ives & Helmus, 2011). More recently, some of the focus have shifted towards
16 approaches that estimate and account for biotic factors, such as competitive or facilitative
17 relationships between species (Ovaskainen *et al.*, 2017). The idea is that by untangling the
18 ways in which such biotic and abiotic factors shape species' distributions, we can gain a
19 mechanistic understanding on how ecological communities are established and change over

time. However, while these factors can increase the predictive performance of some of the models (Norberg *et al.*, 2019), the interpretation of the corresponding parameter estimates has been often questioned (Gotelli & Ulrich, 2010; Harris, 2016; Thurman *et al.*, 2019). This was best illustrated by Blanchet *et al.* (2020), who used basic statistical arguments to highlight the artefactual nature of the link between co-occurrence and species' ecological interactions drawn by some distribution models.

The value of gaining a mechanistic understanding of species' distributions is unquestionable (ref), with several studies highlighting the importance of factors such as biotic interactions and species' dispersal ability in setting their range limits (Wisz *et al.*, 2013; ?; Neuschulz *et al.*, 2018). That said, a lot can be learned from taking a phenomenological approach, focussing instead on the description of basic properties of species' realized niches. For example, the study of species' range sizes along environmental gradients can reveal general biodiversity patterns that are crucial from a conservation and management perspective (e.g. 'Rapoport's rule' Stevens 1992). Differences in species' responses to the environment could shed light on how climatic processes and historical contingencies have shaped their distributions (Rohde, 1992; ?). Other properties, such as the skewness of species' distributions, can also reveal general underlying processes regarding species' physiological tolerance to different environmental conditions (e.g. 'asymmetric abiotic stress limitation hypothesis' ?). More generally, understanding the shape of species' realized niches and the extend to which these vary across species is a crucial issue in ecology and biogeography (ref); however, we do not have an effective way to parsimoniously compare the realized niches of many species. Indeed, there is no general agreement on the shape of species' distributions (ref).

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

50 occupy finite geographic ranges—i.e. their probability distributions have finite variance—
 51 the most conservative statistical approach is to model these as a Gaussian distributions
 52 (?). This is rarely the starting point in most statistical frameworks that study general
 53 biodiversity patterns (but see ref), choosing to use instead Gaussian-logit response curves
 54 (refs). Other factors might then condition species distributions to showcase heavy-tails or a
 55 skewed shapes, revealing interesting ecological processes shaping biodiversity patterns (??).
 56 Second, the aforementioned structural constraints also limit our ability to include any prior
 57 information to our parameter estimates. Observations on species’ geographic variation and
 58 optimal climatic conditions have long been documented, with extensive databases compiled
 59 by botanists and field ecologists documenting basic knowledge on species’ realized niches
 60 (e.g. Landolt *et al.* 2010). That said, this information is rarely accounted for in most
 61 modelling approaches, mainly because there is not a straightforward way to feed this infor-
 62 mation into the parameters of a linear model (Scherrer & Guisan 2019; but see ter Braak &
 63 Looman 1986). Finally, and perhaps most importantly, a direct biological interpretation of
 64 parameter estimates in linear models becomes increasingly difficult as one moves from uni-
 65 modal and symmetric distributions (Jamil & ter Braak, 2013; ter Braak & Looman, 1986)
 66 to skewed distributions (Huisman *et al.*, 1993), making the tests of hypothesis on global
 67 biodiversity patterns particularly challenging. For example, Huisman *et al.* (1993) pro-
 68 posed several non-linear models to characterize several features of species’ response curves;
 69 however, species’ environmental indicator values, range size or distribution skewness are
 70 difficult to capture following these model structures.

71 The field of ecology has quickly moved towards mechanistic and process-based approaches
 72 to understand species’ distributions (Warton *et al.*, 2015). This has resulted in a plethora
 73 of models accounting for several biotic and abiotic factors into the predictions of species
 74 co-occurrence. Here, we instead rethink traditional modelling approaches and develop a
 75 conceptually simple—and yet statistical and computationally complex—statistical frame-
 76 work to revisit some classic hypothesis in ecology and biogeography. In particular, we develop
 77 a Bayesian hierarchical model that accounts for all prior information that we have regarding
 78 the distribution of alpine plant species along an elevation gradient in the Swiss Alps, in-
 79 cluding expert knowledge on species environmental indicator values, range sizes, and plant
 80 physiology. We start by considering species response curves as Gaussian distributed, and

then we adapt such a model to allow for skewed distributions. Using such a statistical framework, we are able to compare the basic properties of the realized niches of multiple species, testing for the existence of general biogeographical patterns. First, we test for the Rapoport's rule, which predicts a positive relationship between range size and elevation (Stevens, 1992). While this pattern has been largely studied for multiple systems and across gradients (McCain & Knight, 2013); contrasting evidence suggests this rule not to be pervasive across species (McCain & Knight, 2013). Moreover, we test the extend to which different types of species, such as native or neophytes, present variation in this geographical pattern. Second, we study whether or not species' distributions show steeper declines towards stressful conditions, testing the so-called abiotic stress limitation hypothesis (ref). We tested this for vegetation data using (Huisman *et al.*, 1993)'s statistical models

Moreover, this process allows us to understand the value of our prior information for explaining the observed patterns as well as to detect unexplained variation in the distribution of species.

In this study, ...

regarding the way systems of many species are distributed along an environmental gradient. In particular, we use models that is solely constrained by the empirical information that we truly have regarding a particular system, relaxing as much as possible the structural constraints of the statistical framework. Then, we want to use this framework to answer basic aspects regarding the way systems of many species are distributed along an environmental gradient.

revisit two classical hypothesis in ecology and biogeography. This pattern has been largely studied for multiple systems and across gradients (McCain & Knight, 2013); however, contrasting observations .

including the limited ability to include prior information to inform parameter estimates, somewhat biologically meaningless interpretation of these parameters, and poor out of sample fits of occurrence data. From a purely statistical perspective, we know that species occupy a geographic range; therefore, we know that their distributions have finite variance. Indeed, observations on species' geographic variation and optimal climatic conditions have been long documented, with extensive databases compiled by botanists and field ecologists

111 documenting basic knowledge on species' distributions. While other factors might influence
112 species' co-occurrence (e.g. the influence of the aforementioned biotic interactions among
113 species; ref), we do not necessarily have an intuition of how exactly these factors will
114 influence the shape of species' distributions. As a result, the most conservative assumption
115 regarding species' response curves and the safest bet—i.e. the one with the largest entropy—
116 is that these follow a Gaussian distribution.

117 Under this premise, we develop a statistical framework to ...

118 In practice, distribution frameworks often assume an underlying linear relationship be-
119 tween covariates (but see XX and YY). This is useful because ... However, it results in
120 many statistical issues: (i) no prior information added; (ii) meaningless parameters; and
121 (iii) poor out of sample fits.

122 From a purely statistical perspective, we know that species occupy a geographic range;
123 we know that their distributions have finite variance. Indeed, observations on species'
124 geographic variation and optimal climatic conditions have been long documented, with
125 extensive databases compiled by botanists and field ecologists documenting basic knowledge
126 on species' distributions. Therefore, the most conservative assumption regarding species'
127 response curves and the safest bet—i.e. the one with the largest entropy—is that these
128 follow a Gaussian distribution.

129 To decide among modelling approaches, we first need to agree on what we know about
130 the system. We know that species occupy a geographic range; therefore, we know that
131 their distributions have finite variance. Indeed, observations on species' geographic varia-
132 tion and optimal climatic conditions have been long documented, with extensive databases
133 compiled by botanists and field ecologists documenting basic knowledge on species' dis-
134 tributions. One could point out that we also know that many other factors might influ-
135 ence species' presence/absence—e.g. the influence of the aforementioned biotic interactions
136 among species. However, we do not necessarily have an intuition of how exactly these fac-
137 tors will influence the shape of species' distributions. Therefore, if all we truly knew about
138 a species' distribution was that they have finite variance, the most conservative assumption
139 and the safest bet—i.e. the one with the largest entropy—is that such distribution is a
140 Gaussian.

141 For example, species' physiological tolerance to climatic conditions.

142 There is not an easy way to untangle the true shape of species' distributions, as this
143 shape is likely to showcase idiosyncrasies at the species level and across systems. The
144 aim of this work, it is not to answer these questions nor to provide a general approach
145 that accommodates such idiosyncrasies. Instead, we want to use a model that is solely
146 constrained by the empirical information that we truly have regarding a particular system,
147 relaxing as much as possible the structural constraints of the statistical framework. Then,
148 we want to use this model to answer basic aspects regarding the way systems of many
149 species are distributed along an environmental gradient.

150 To decide among modelling approaches, we first need to agree on what we know about
151 the system. We know that species occupy a geographic range; therefore, we know that their
152 distributions have finite variance. Indeed, observations on species' geographic variation and
153 optimal climatic conditions have been long documented, with extensive databases compiled
154 by botanists and field ecologists documenting basic knowledge on species' distributions.
155 One could point out that we also know that many other factors might influence species'
156 presence/absence—e.g. the influence of biotic interactions among species. However, we do
157 not necessarily have an intuition of how exactly these factors will influence the shape of
158 species' distributions. As a result, if all we truly knew about a species' distribution was
159 that they have finite variance, the most conservative assumption and the safest bet—i.e.
160 the one with the largest entropy—is that such distribution is a Gaussian.

161 Unfortunately, while some of these approaches can increase the predictive performance
162 of distribution models (ref), the nature of some of the estimates have been theoretically
163 questioned (ref). For example...

164 Increasing efforts have been devoted to improving the ability of statistical models to
165 predict the presence/absence of species across ranges (ref). Accurately predicting how
166 species are distributed across ranges is crucial for understanding the impacts and effect of
167 global climate change. With that said, accounting for the mechanisms can come at the
168 cost of modelling noise... However, much less attention is paid to how species' distributions
169 compare to each other.

170 Increasing efforts have been devoted to improving the ability of statistical models to

171 predict the presence/absence of species across ranges. However, much less attention is paid
172 to how species' distributions compare to each other.

173 While the value of gaining a mechanistic understanding of species' distributions is un-
174 questioned, a lot can be learned from shedding light on basic properties of species' realized
175 niches.

176 Accounting for expert knowledge on species' environmental preferences to understand
177 general distribution patterns.

178 Here, we...Bayesian framework... This allow us to account for... as well as to tackle
179 long-standing hypothesis regarding basic aspects of species distributions.

180 For non-parametric models, these often do not account for species uncertainties. One
181 can benefit a lot from a bayesian framework. First, it allows us to compare the parameter
182 estimates ... It also help us incorporate prior information.

183 In this work, we first....

184 There is no general agreement the shape of species distributions. While many ecological
185 textbooks (Begon et al., 1990, Giller, 1984, Krebs, 1994) assume this to be unimodal and
186 symmetric, some have warned that empirical distributions can take many different forms
187 (?). There is not an easy way to untangle the true shape of species' distributions, as this
188 shape is likely to showcase idiosyncrasies at the species level and across systems. The
189 aim of this work, it is not to answer these questions nor to provide a general approach
190 that accommodates such idiosyncrasies. Instead, we want to use a model that is solely
191 constrained by the empirical information that we truly have regarding a particular system,
192 relaxing as much as possible the structural constrains of the statistical framework. Then,
193 we want to use this model to answer basic aspects regarding the way systems of many
194 species are distributed along an environmental gradient.

195 To decide among modelling approaches, we first need to agree on what we know about
196 the system. We know that species occupy a geographic range; therefore, we know that their
197 distributions have finite variance. Indeed, observations on species' geographic variation and
198 optimal climatic conditions have been long documented, with extensive databases compiled
199 by botanists and field ecologists documenting basic knowledge on species' distributions.

One could point out that we also know that many other factors might influence species' presence/absence—e.g. the influence of biotic interactions among species. However, we do not necessarily have an intuition of how exactly these factors will influence the shape of species' distributions. As a result, if all we truly knew about a species' distribution was that they have finite variance, the most conservative assumption and the safest bet—i.e. the one with the largest entropy—is that such distribution is a Gaussian.

Scarce data and little to no attempt to account for uncertainty in the predictions. Similar to rapoport's rule, we can also ask other questions regarding general geographical patterns of species distributions.

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 studied 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, following 30 years (1961–1990) of meteorological data from national weather stations, Scherrer & Guisan (2019) calculated multiple climatic variables for each site at high spatial resolution (25 m). Here, we focussed on 9 climatic variables, including: 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.

227 *Floristic data*

228 To complement the aforementioned distribution data, we used a floristic database of most
229 vegetation across Switzerland. This database was build based on expert knowledge and
230 field experience of botanists and ecologists, and contains information regarding species’
231 environmental preferences and physiological traits. Species’ environmental preferences in
232 this database can be used to inform distribution models—e.g. as an informative prior in
233 a Bayesian framework. These are characterized following the ecological indicator values
234 developed by Landolt *et al.* (2010), providing both an estimate of the average conditions in
235 which a species can be found and a broad description of their range of variation. These val-
236 ues are provided for a range of 10 climatic variables, including temperature, continentality,
237 light conditions, as well as moisture, acidity and nutrient content of the soil (see a full list
238 and description of the ecological indicators in the Supplementary Methods; Landolt *et al.*
239 2010). On the other hand, the information regarding species’ physiological traits represent
240 general descriptions of species’ growth and life strategies—examples include their growth
241 forms, nature of the storage organs, dispersal ability and pollinator agents. In total, we
242 identify more than 120 binary traits that characterize the physiology of species (see a full
243 list and description of the ecological indicators in the Supplementary Methods; Landolt
244 *et al.* 2010).

245 *[Trait data]*

246 This could be Tom’s data if we end up using it.

247 **Distribution model**

248 There is a long list of model structures well suited to characterize species’ distributions (see
249 XX for a review); however, we were interested in a model that explicitly incorporates all in-
250 formation regarding plant’s environmental preferences found in the floristic database. More
251 specifically, we wanted to account for the climatic indicator values and range of variation
252 registered for all plants in our dataset. These two values provide basic information regard-
253 ing plant’s optimal environmental conditions and width of their distributions. Therefore,

we first formulated a baseline model that directly accounts for such prior information.

Baseline model

Given y_{ij} the presence/absence of any species i in any given site j , and a set of k environmental variables x_{jk} , we estimate species' distributions as:

$$\begin{aligned}
y_{ij} &\sim \text{Binomial}(1, p_{ij}) \\
\log(p_{ij}) &= -\alpha_i - \sum_k \lambda_{ik} (x_{jk} - \beta_{ik})^2 \\
\log(\alpha) &\sim \text{MVNormal}(\hat{\alpha}, \Sigma^\alpha) \\
\beta_{ik} &\sim \text{MVNormal}(\hat{\beta}_k, \Sigma^{\beta_k}) \\
\log(\lambda_{ik}) &\sim \text{MVNormal}(\hat{\lambda}_k, \Sigma^{\lambda_k}) \\
\hat{\alpha}, \hat{\lambda}^k, \hat{\beta}^k &\sim \text{Normal}(0, 1)
\end{aligned} \tag{1}$$

Notice that this model structure assumes all plants to have a uni-modal distributions along each environmental axis (see the model's behaviour in Supplementary Figure XX), where parameters α_i , β_i^k , and λ_i^k describe amplitude of the probability p_{ij} , species' average climatic suitability and range of variation along the different environmental gradients, respectively[†]. While potentially sacrificing predictive accuracy, this model structure allows us to explicitly incorporate all prior knowledge that we have regarding species' distributions via Σ^α , Σ^{β_k} and Σ^{λ_k} . More specifically, we express β_i^k and $\log(\lambda_i^k)$ as multivariate normal distributions—i.e. Gaussian processes—such that Σ^{β_k} and Σ^{λ_k} are variance-covariance matrices describing species' similarity in terms of their average climatic suitability and range of variation along the different environmental gradients, respectively. Likewise, $\log(\alpha)$ is characterized as a Gaussian Process, where the corresponding variance-covariance matrix Σ^α is designed to also incorporate some of the prior information that we have with regards to species' physiological traits.

[†]I'll rewrite the likelihood function to an ordered categorical as soon as I get things to work properly with count data.

271 In all cases, all variance-covariance matrices are defined as follows:

$$\Sigma_{ij}^{\chi} = \eta_{\chi} \exp \left(-\rho_{\chi} D_{ij}^{\chi 2} \right) + \delta_{ij} \sigma_{\chi}, \quad (2)$$

272 where Σ_{ij}^{χ} describes the covariance between any pair of species i and j for any given
 273 parameter α_i , β_i^k , and λ_i^k . Following this expression, such covariance declines exponentially
 274 with the square of the different D_{ij}^{χ} , which are distance measures computed using the
 275 prior information that we have regarding species' distributions. Specifically, given α_i , β_i^k ,
 276 and λ_i^k , the distance measures are calculated using plants' physiological traits, ecological
 277 indicator values and range of variation, respectively (see below for further details). For each
 278 covariance matrix, the hyperparameter ρ_{χ} determines the rate of decline of the covariance
 279 between any two species, and η_{χ} defines its maximum value. The hyperparameter σ_{χ}
 280 describes the additional covariance between the different observations for any given species.
 281 For any given hyperparameter, we choose adaptive priors across covariance structures.
 282 That is, and taking ρ_{χ} as an example, we choose a prior $\log(\rho_{\chi}) \sim \text{Normal}(\hat{\rho}, \sigma_{\rho})$ such that
 283 $\hat{\rho} \sim \text{Normal}(0, 1)$ and $\sigma_{\rho} \sim \text{Exponential}(1)$. Similar priors were chosen for both η_{χ} and
 284 σ_{χ} . We generated the posterior samples for the Bayesian models with the help of the R
 285 package 'rstan' to (?).

286 *Distance matrices*

287 The missing component in the description of model (1) is the distance matrices D^{χ} used
 288 to define the covariance matrices Σ^{α} , Σ^{β_k} and Σ^{λ_k} . In this model, such distance matrices
 289 characterize differences between plant species. In the floristic data, however, the prior infor-
 290 mation that we have for these differences is represented by a set of ordinal and categorical
 291 traits. More specifically, both the ecological indicator values and range of variation—which
 292 define the prior information that we have for β_i^k , and λ_i^k , respectively—are ordinal traits
 293 specified for all species. In contrast, the plants' physiological data—shaping the prior for
 294 the parameters α_i —are characterized by categorical data containing multiple missing en-
 295 tries. Therefore, we need to carefully compile this data into distance matrices in order to
 296 be able to feed this prior information into the model.

More generally, we want to understand the way N species are characterized by M categorical traits. One way to frame this problem is by using a network representation. Following the ideas presented by Godoy-Lorite *et al.* (2016), we assume that species can be connected to each of these traits by an interaction (i, j) that can be of any type $r \in R$. Notice that this provides as with multiple ways to account for the information—and lack thereof—contained in the different categorical and ordinal traits M . That is, the R types of interactions can represent the lack of information for a particular link (i, j) , the absence or presence of such interaction, and any type of association between i and j .

Given a set of interactions R^* between N and M , we use a Mixed Membership Stochastic Block Model (MMSBM) to characterize these. In particular, we consider that plants and traits can be classified into K and L groups, respectively. For every species i , we assume that there is a probability $\theta_{i\alpha}$ for it to belong to any of the K species groups. Likewise, we also assume that any trait j has a probability $\phi_{j\beta}$ of belonging to any of the L trait groups. Finally, we define $p_{\alpha\beta}(r)$ as the probability of a species from group α interacting with a trait from group β by an association type r . Putting these together, the probability of an interaction (i, j) of type r can be calculated as:

$$Pr[r_{ij} = r] = \sum_{\alpha\beta} \theta_{i\alpha} \phi_{j\beta} p_{\alpha\beta}(r) \quad (3)$$

Following this definition, we want to find the group memberships that maximize the likelihood $P(R^*|\theta, \phi, p)$. Doing so is difficult optimization problem; however, it has been shown that one can estimate the different $\theta_{i\alpha}$, $\phi_{j\beta}$, and $p_{\alpha\beta}(r)$ parameters by maximizing the likelihood using an expectation-maximization algorithm (Godoy-Lorite *et al.*, 2016; Tarrés-Deulofeu *et al.*, 2019). In simple terms, one can iteratively find multiple local minima for the likelihood, and average over the estimated the parameter values (Godoy-Lorite *et al.*, 2016)[†].

[†]While this averaging is trivial for the estimated probabilities $Pr[r_{ij} = r]$, it is non-trivial if one wants to find averages for the group memberships. The reason for this is related to the stochastic nature of the expectation-maximization algorithm. This algorithm initially assigns random group memberships to both species and traits. While this random labelling is irrelevant when studying the probabilities $Pr[r_{ij} = r]$, it is instead crucial for averaging $\theta_{i\alpha}$, $\phi_{j\beta}$, and $p_{\alpha\beta}(r)$. Therefore, before averaging the group membership estimates, one needs to find the bijective relationship for the labellings of different iterations of the optimization algorithm. In a nutshell, for every iteration, I do this by using a simulated annealing algorithm on the estimated $p_{\alpha\beta}(r)$, matching the corresponding labelling to a reference iteration.

320 The average estimates for the group memberships provide us with a different scale to
 321 classify species based on the traits these have. In short, for any species i , we can esti-
 322 mate a K -dimensional vector $\vec{\theta}_i$ that describes the extend to which i belong to each group
 323 membership—i.e. the extend to which a species is of one type or another. This classification
 324 is useful because it can be used to compare species, defining a way to measure the distance
 325 between species based on an arbitrary—and potentially incomplete—set of categorical or
 326 ordinal traits M . The simplest case is to define the distance as $D_{ij} = |\vec{\theta}_i - \vec{\theta}_j|$. Alterna-
 327 tively, one could also define K distance matrices based on the different group memberships
 328 $D_{ij}^\alpha = |\theta_{i\alpha} - \theta_{j\alpha}|$.

329 *Modifying the variance-covariance structures*

330 The model structure defined in Eq. (1) allows us to test the effect of adding new information.
 331 Specifically, we can do this by modifying Eq. (2). For example, imagine that we have
 332 multiple matrices D^k characterizing species' differences along different axis of variation—
 333 i.e. two matrices characterizing ecological and environmental traits, or multiple matrices
 334 resulting from the different group memberships estimated using the MMSBM. One could
 335 modify Eq. (2) for a particular parameter—e.g. parameter α_i —such that

$$\Sigma_{ij}^\alpha = \eta_\alpha \exp \left(- \sum_k \rho_{\alpha k} D_{ij}^{k,2} \right) + \delta_{ij} \sigma_\alpha, \quad (4)$$

336 where now $\rho_{\alpha k}$ are separate relevance hyperparameters for each distance matrix in the total
 337 variance of α_i . Notice that the same is true for the covariance of parameters β_i^k and λ_i^k .
 338 Finally, for all hyperparameters and as described for the baseline model, we use adaptive
 339 priors across covariance structures.

340 **Results**

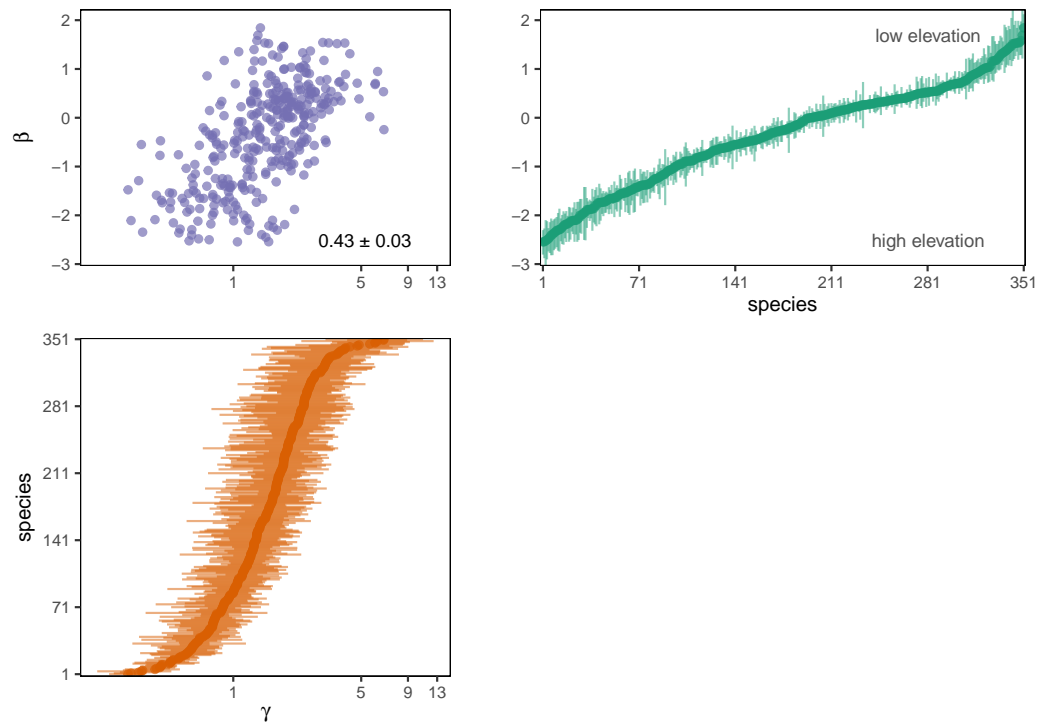


Figure 1: Relationship between mean and variance of species' distributions. These are the results for the main axis of variation for the climatic data (results for the second axis of variation presented in the Supplementary Fig. 2).

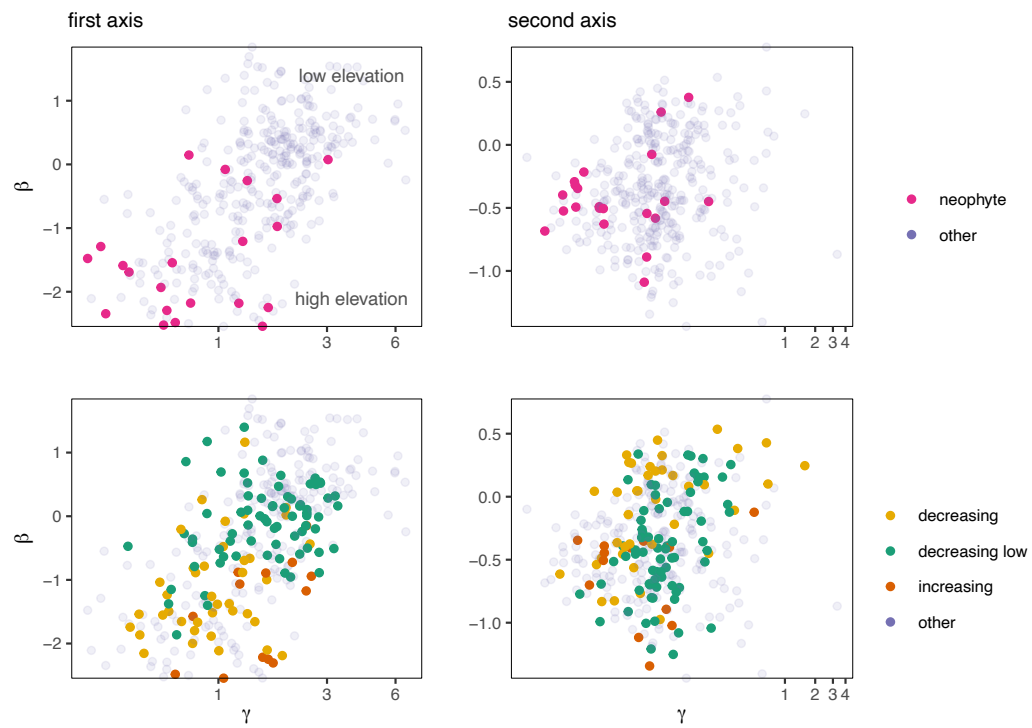


Figure 2: Are there clear geographical patterns for neophytes and for species with decreasing or increasing abundance?

References

- Austin, M. P. (1987). Models for the analysis of species' response to environmental gradients. *Vegetatio*, 69, 35–45.
- Blanchet, F. G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, 23, 1050–1063.
- 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.
- 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.
- 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.
- McCain, C. M. & Knight, K. B. (2013). Elevational Rapoport's rule is not pervasive on mountains. *Global Ecology and Biogeography*, 22, 750–759.

366 Neuschulz, E. L., Merges, D., Bollmann, K., Gugerli, F. & Böhning-Gaese, K. (2018).
367 Biotic interactions and seed deposition rather than abiotic factors determine recruitment
368 at elevational range limits of an alpine tree. *Journal of Ecology*, 106, 948–959.

369 Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., Araújo,
370 M. B., Dallas, T., Dunson, D., Elith, J., Foster, S. D., Fox, R., Franklin, J., Godsoe, W.,
371 Guisan, A., O’Hara, B., Hill, N. A., Holt, R. D., Hui, F. K. C., Husby, M., Kålås, J. A.,
372 Lehtikainen, A., Luoto, M., Mod, H. K., Newell, G., Renner, I., Roslin, T., Soininen, J.,
373 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 Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F. G., Duan, L., Dunson, D., Roslin,
378 T. & Abrego, N. (2017). How to make more out of community data? A conceptual
379 framework and its implementation as models and software. *Ecology Letters*, 20, 561–576.

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

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

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

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

388 Tarrés-Deulofeu, M., Godoy-Lorite, A., Guimerà, R. & Sales-Pardo, M. (2019). Tensorial
389 and bipartite block models for link prediction in layered networks and temporal networks.
390 *Physical Review E*, 99, 032307.

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

- 393 Thurman, L. L., Barner, A. K., Garcia, T. S. & Chestnut, T. (2019). Testing the link
394 between species interactions and species co-occurrence in a trophic network. *Ecography*,
395 42, 1658–1670.
- 396 Warton, D. I., Blanchet, F. G., O’Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C.
397 & Hui, F. K. C. (2015). So Many Variables: Joint Modeling in Community Ecology.
398 *Trends in Ecology & Evolution*, 30, 766–779.
- 399 Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dor-
400 mann, C. F., Forchhammer, M. C., Grytnes, J.-A., Guisan, A., Heikkinen, R. K., Høye,
401 T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.-C., Normand, S., Öckinger, E.,
402 Schmidt, N. M., Termansen, M., Timmermann, A., Wardle, D. A., Aastrup, P. & Sven-
403 ning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised
404 assemblages of species: Implications for species distribution modelling. *Biological Re-*
405 *views*, 88, 15–30.