

Model linearity breeds contempt: using Bayesian non-linear models to uncover broad macroecological patterns

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

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

1 Abstract

2 Species' realized niches are classically pictured as bell-shaped probability distributions. These
3 distributions, however, can actually take many different forms. For example, fat-tailed or
4 skewed responses are very common across fields, as these can naturally emerge as a result of
5 several ecological processes. While one does not need to know the shape of species' distribu-
6 tions to effectively model them, studying their basic form can teach us a lot about the ways
7 climatic processes and historical contingencies have shaped ecological communities. Unfor-
8 tunately, we still lack a general understanding of the basic properties describing the shape of
9 species' distributions, and much less is known about how these compare to each other across
10 gradients. Here, we use a set of Bayesian non-linear models to uncover such properties.
11 These models account for all prior knowledge we have regarding species' realized niches, in-
12 cluding expert knowledge of their environmental preferences and ecological strategies. With
13 this approach, we are able to distil the shape of empirical plant distributions, which helps
14 us tackle long-standing hypotheses regarding the way ecological communities are assembled
15 across space. In particular, we studied the relationship between several properties of distri-
16 butions, such as the link between species' range size and elevation, revealing the existence
17 of broad macroecological patterns along environmental gradients. Moreover, we are able to
18 shed light on the extent to which some aspects of the shape of observed realized niches—
19 such as kurtosis and skewness of the distributions—could be intrinsic properties of species'
20 historical contexts. Overall, our approach offers a useful statistical framework to understand
21 the shape of species' distributions, and our results provide an unprecedented perspective of
22 the way systems of many species are distributed along environmental gradients.

23 Introduction

24 One of the central goals of ecology is to understand the ways species are distributed across
25 space and time. While many ecological textbooks assume the shape of species' realized
26 niches to be unimodal and symmetric along environmental gradients (Krebs, 1972), some
27 have warned that empirical distributions can take many different forms (Austin, 1987, 2002;
28 ?). In practice, there is a strong argument to be made in favour of assuming these to be
29 bell shaped. Namely, if all that we are willing to assume about species' distributions is that
30 these occupy finite geographic ranges, the most conservative statistical approach is to model
31 their distribution as Gaussian (i.e. the corresponding maximum entropy distribution; Frank
32 2009). That said, there is currently no general agreement on the basic shape of species'
33 realized niches. Indeed, many factors can play a role in defining their shape, and several
34 natural processes can lead to non-normal distributions.

35 Fat-tailed and skewed distributions are very common across fields. The former natu-
36 rally emerges as a result of processes involving seasonality (e.g. in communications patterns;
37 Malmgren *et al.* 2008) or some stochastic events (e.g. in the spread of infectious diseases;
38 Wong & Collins 2020). Indeed, species' dispersal patterns have been shown to have fat tails
39 due to the natural variability among individuals (Petrovskii *et al.*, 2009). This is important
40 because one might expect environmental and individual variation to also be crucial factors
41 determining the presence and absence of species along gradients, and fat tails are therefore a
42 plausible property of species' realized niches. Similarly, several processes can lead to skewed
43 distributions. For example, species might present asymmetric environmental tolerances along
44 altitudinal gradients, allowing them to withstand different temperature extremes (Sunday
45 *et al.*, 2011). Species might also experience abiotic and biotic pressures that increase or de-
46 crease along a temperature gradient, which could result in species' distributions presenting
47 steeper declines towards warmer or colder environments (Normand *et al.*, 2009). Overall,
48 many different properties could characterize species' realized niches, and every new shape
49 entails different underlying hypotheses regarding the way communities are assembled over
50 time (D'Amen *et al.*, 2017).

51 Comparing these properties across species allows us to study broad macroecological pat-
52 terns that could be critical from a conservation and management perspective (Stevens, 1992;

53 ?). For example, the Rapoport's rule, a classic biogeographical hypothesis, predicts species'
54 ranges to increase with latitude or elevation (Stevens, 1992; ?), hinting at the existence of
55 general biogeographical constraints that shape species' distributions along gradients. This
56 sort of macroecological patterns are interesting because they provide insights into the way
57 different species assemble and establish in different environments. That is, the differences in
58 species' responses to the environment can shed light on how climatic processes and historical
59 contingencies have differently shaped their distributions (Rohde, 1992; ?). Uncovering the
60 shape of species' realized niches and the extent to which these vary across species is never-
61 theless a challenging statistical problem to solve. Indeed, to this date, we do not have an
62 effective way to parsimoniously compare the shape of the realized niches of many species
63 along environmental gradients.

64 Over the last two decades, ecologists have developed a plethora of distribution models to
65 try to untangle the factors that play a role in defining species' realized niches (Guisan & Zim-
66 mermann, 2000). These models are fundamental to the scientific community for predicting
67 changes in species' geographic distributions and the effects of environmental disturbances.
68 Such frameworks, however, commonly assume an underlying linear relationship between co-
69 variates (but see 'semiparametric models'; Norberg *et al.* 2019). This is useful because it
70 simplifies the optimization process but might not be ideal when studying and comparing the
71 shape of species' distributions along environmental gradients. First and foremost, a linear
72 relationship between covariates often comes with a set of implicit mathematical constraints
73 that might not be biologically justified. While this might not hinder the predictive perfor-
74 mance of the models (Norberg *et al.*, 2019), a direct biological interpretation of parameter
75 estimates in linear models becomes increasingly difficult as one moves from unimodal and
76 symmetric distributions (ter Braak & Looman, 1986; Jamil & ter Braak, 2013) to fat-tailed
77 or skewed responses (Huisman *et al.*, 1993). Second, the aforementioned structural con-
78 straints also limit our ability to include any prior information to our parameter estimates.
79 Observations on species' geographic variation and optimal climatic conditions have long been
80 documented, with extensive databases compiled by botanists and field ecologists document-
81 ing basic knowledge of species' realized niches (e.g. Landolt *et al.* 2010). That said, this
82 information is rarely accounted for in most modelling approaches, likely because there is not
83 a straightforward way to feed this information into the parameters of a linear model (Scher-

rer & Guisan 2019; but see ter Braak & Looman 1986; Ovaskainen *et al.* 2017). Finally,
Huisman *et al.* (1993) proposed several non-linear structures to characterize several features
of individual species' response curves. Setting aside the fact that the interpretation and
comparison of parameter estimates becomes challenging following these model structures,
these were not designed to jointly study different species, taking full advantage of modern
statistical approaches (e.g. sharing information among species or accounting for parameter
uncertainty; Evans *et al.* 2016).

In this work, 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
the distribution of plant species along an elevation gradient in the Swiss Alps, including
expert knowledge of species environmental indicator values, range sizes, and plant ecological
strategies. We start by considering species' response curves as Gaussian distributed, and
then we adapt our model to allow non-linear responses characterizing skewed and long-tailed
distributions. Using this statistical framework, we are able to compare the basic properties of
the realized niche of multiple species, testing for the existence of broad macroecological patterns.
Comparing the posterior distribution of those parameters that control for the shape
of distributions, we are also able to showcase variation in the way different types of species,
such as native or neophytes, might respond to the environment. More generally, we are able
to uncover the approximate shape of empirical plant distributions and answer fundamental
questions regarding the way systems of many species are distributed along environmental
gradients.

Methods

Empirical data

We studied the distribution of 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 around 5500 vascular plants across Switzerland. Some of the information in this database has been previously shown to account for unexplained variation when used as explanatory variables in species' distribution models (Scherrer & Guisan, 2019). It was built based on expert knowledge and phytosociological field experience of botanists and ecologists, and contains information regarding plants' environmental preferences and ecological strategies.

Species' environmental preferences in this database can be used to inform distribution models—e.g. as an informative prior in a Bayesian framework. These are characterized following the ecological indicator values developed by Landolt *et al.* (2010), providing both an estimate of the average conditions in which a species can be found as well as a broad description of their range of variation. These values are provided for a range of 10 environmental variables, including temperature, continentality, light conditions, as well as moisture, acidity

and nutrient content of the soil (see a full list and description of the ecological indicators in the Supplementary Table 1; Landolt *et al.* 2010). On the other hand, the information regarding species' ecological strategies represents general descriptions of species' growth and life strategies—examples include their growth forms, nature of the storage organs, dispersal ability and pollinator agents. In total, we identify more than 120 binary traits that characterize the physiology and life-history of species (see a full list and description of the ecological indicators in the Supplementary Table 1; Landolt *et al.* 2010). Finally, and in addition to species' environmental preferences and physiological traits, the floristic data also contains information on species introduction status (e.g. identifying those species that are neophytes or archaeophytes) and change tendency (e.g. indicating species that have shown decline or increase in their populations over the recent decades). We describe this information in more detail in the Supplementary Table 1.

Baseline model

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

Response curve

To choose an appropriate response curve, we first need to agree on what we truly know about the system. Given the prior information that we have about the system, we know that species occupy specific geographic ranges; therefore, we know that their distributions have finite variance. While we could also assume that many other factors might influence species' presence in a given site—e.g. the biotic interactions among species in the site—we do not necessarily have an *a priori* expectation of how exactly these factors will influence the

168 shape of species' distributions. Therefore, for this baseline model, if all that we are willing to
 169 assume about species' realized niches is that these have finite variance, the most conservative
 170 assumption and the safest bet—i.e. the one with the largest entropy—is that they follow a
 171 Gaussian distribution (Fig 1a). That is, given the presence/absence or abundance y_{ij} of
 172 any species i in any given site j , and an environmental variable x_j , we can define species'
 173 responses to the environment as

$$y_{ij} \sim F(p_{ij})$$

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

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

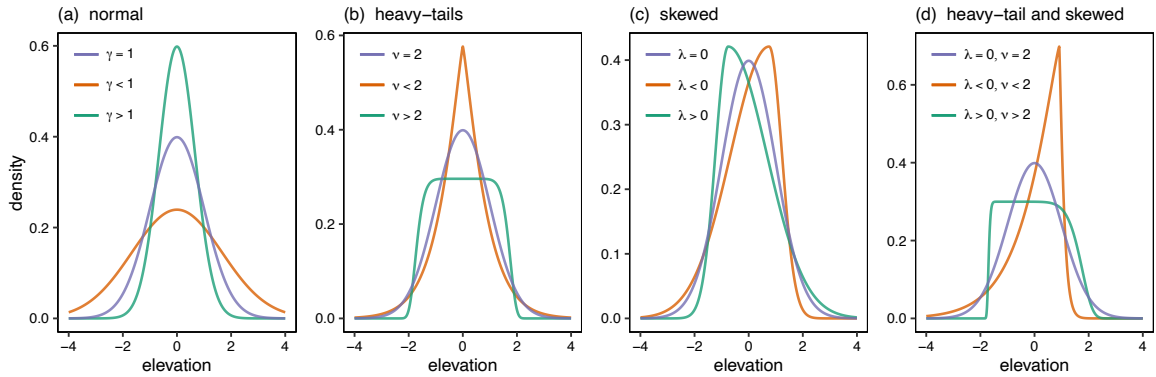


Figure 1: Different response curves. Panel (a) shows the probability density function characterized by Eq. (1) for different values of γ , when $\mu = 0$. Panel (b) shows the probability density function characterized by Eq. (5) for different values of ν , when $\gamma = 1$ and $\mu = 0$. Panel (c) shows the probability density function characterized by Eq. (6) for different values of λ , when $\gamma = 1$ and $\mu = 0$. Panel (d) shows the probability density function characterized by Eq. (7) for different values of λ and ν , when $\gamma = 1$ and $\mu = 0$. Notice that for each case, we chose α values that normalize the probability distributions.

183 *Model priors*

184 The model structure described above allows us to explicitly incorporate all prior knowledge
 185 that we have regarding species' distributions contained in the floristic database. To do so,
 186 we define the prior distributions for the parameters in model (1) as:

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

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

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

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

Alternative variance-covariance structures

The model structure defined above allows us to test how different sources of information 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 packages 'rstan' and 'cmdstanr' (Stan Development Team, 2021). Sampling models like the ones described above can be computationally very intensive. 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 at least 20 occurrences when modelling both binary data and ordinal data.

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). To perform the data analysis and generate the figures, we used some of the functions available with the R package 'rethinking' (?).

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 prior information. We focused on two properties: fat-tailed and skewed responses.

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 shapes. We focused on two properties: fat-tailed and skewed responses. While there are several model structures that could account for these properties, we propose new species' response curves following three criteria. First, the probability distribution of a species along an environmental gradient must have a defined mean and variance. This is important because we know that species naturally have different environmental preferences as well as finite geographic ranges. Second, the Gaussian shape must be a special case of the probability distribution, allowing species to showcase variation regarding the presence (or lack thereof) of any given pattern. Finally, there must be a re-parametrization of the model that allows us to keep the same prior information and interpretable parameters.

Fat-tailed response curve

Fat-tailed distributions represent distributions with relatively high representation of extreme events. While many different distributions exhibit this property, we decided to accommodate this feature into our baseline model by considering a response curve that follows a generalized error distribution. Such a distribution is useful because the Gaussian shape is a special case of it, and it contains a parameter that regulates the level of kurtosis—ranging from longer to shorter tails than the Gaussian case (Fig 1b). In particular, we can adapt Eq. (1) to present this non-linear form as follows:

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

where $\gamma'_i = g(\gamma_i, \nu_i)$, and ν_i is a parameter that describes the kurtosis of the distribution, which we define as $\nu_i \in (1, \infty)$. Following this, we choose an adaptive prior for this set 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 $\sigma_\nu \sim \text{Exponential}(2)$. Given the relationship between γ'_i and γ_i , we can re-parametrize the

model and follow Eq. (2) to define the prior distributions (see Supplementary Table 2; Code Availability section; Nadarajah 2005). Notice that the Gaussian distribution will naturally emerge when $\nu_i = 2$.

Alternatively, we could have used other distributions that present fat tails and fulfil the selection criteria described above. For example, the non-standardized Student's t-distributions is an interesting distribution because, as opposed to the generalized error distribution, it allows for fat tails without generating a cusp at the center (see Fig 1b). However, we avoided using the non-standardized Student's t-distributions because it does not allow for tails that are lighter than normal (e.g. $\nu_i > 2$ in Eq. 5; Fig 1b), and the sampling of the model can be somewhat more challenging (ref).

Skewed response curve

Skewed responses present steeper declines towards either side of the distribution. One way to accommodate this feature in our models is by considering a skewed normal distribution (Supplementary Figure XX). As for the case described above, the Gaussian is a special case of this distribution, and it contains a parameter that controls for the level and direction of 'skewness' (Fig 1c). Importantly, this distribution presents normal-like tails; therefore, the added skewness does not make additional assumptions regarding how species are distributed along the gradient. To test for the existence of this feature, we modified Eq. (1) as

$$\log(p_{ij}) = -\alpha_i - \gamma'_i \left(\frac{x_j - \beta'_i}{1 + \lambda_i \operatorname{sgn}(x_j - \beta'_i)} \right)^2, \quad (6)$$

where $\gamma'_i = q_1(\gamma_i, \nu_i, \lambda_i)$, $\beta'_i = q_2(\gamma_i, \beta_i, \nu_i, \lambda_i)$, and λ_i is a parameter that describes the skewness of the distribution such that $\lambda_i \in (-1, 1)$. The function $\operatorname{sgn}(x)$ characterizes the sign function. We chose λ_i to have an adaptive prior such that $\operatorname{logit}\left(\frac{\lambda_i + 1}{2}\right) \sim \operatorname{Normal}\left(\hat{\lambda}, \sigma_\lambda\right)$, where $\hat{\lambda} \sim \operatorname{Normal}(0, 1)$ and $\sigma_\lambda \sim \operatorname{Exponential}(1)$. Notice that this model can be reparametrized following q_1 and q_2 , allowing us to set the rest of the prior distributions as described for the baseline model (see Supplementary Table 2; Code Availability section). In this case, the Gaussian distribution is a special case of Eq. (6) when $\lambda_i = 0$ (Ashour & Abdel-hameed, 2010).

Finally, one could consider a response curve with both kurtosis and skewness. A convenient way to achieve this is by using a response curve that follows a skewed generalized error distribution. This is a combination of the two distributions described above, containing two parameters that control for both the level and direction of kurtosis and skewness (Fig 1d). The skewed generalized error distribution can be considered by modifying the species' response curve in Eq. (1) as

$$\log(p_{ij}) = -\alpha_i - \left(\frac{\gamma'_i |x_j - \beta'_i|}{1 + \lambda_i \operatorname{sgn}(x_j - \beta'_i)} \right)^{\nu_i}, \quad (7)$$

where $\gamma'_i = f_1(\gamma_i, \nu_i, \lambda_i)$, $\beta'_i = f_2(\gamma_i, \beta_i, \nu_i, \lambda_i)$, and ν_i and λ_i are parameters that control the kurtosis and skewness of the distribution, respectively. We define ν_i , λ_i and their prior distributions as in Eq. 5 and 6, respectively. Again, we can re-parametrize the model following f_1 and f_2 , and set the rest of the prior distributions as in the baseline model (see Supplementary Table 2; Code Availability section). Notice that the generalized error distribution (Eq. 5) and the skew normal distribution (Eq. 6) are special cases of Eq. (7) when $\lambda_i = 0$ and $\nu_i = 2$, respectively.

One way to test the extent to which species' physiological traits inform the kurtosis or skewness of the distribution of any species i is by characterizing ν_i or λ_i as Gaussian processes. That is, we can consider the prior distribution for these parameters as multivariate normal with a variance covariance matrix Σ . As described before, this variance-covariance matrix can be estimated 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 along the main axis of variation of all environmental variables. Using the presence and absence of species across sites as the response variable, we sampled the posterior distributions of the baseline model,

308 accounting for the information in the floristic database regarding species' indicator values
 309 and range of variation. This allowed us to map the center and variance of species' distri-
 310 butions along the environmental gradient (Fig. 2). Studying the relationship between these
 311 properties, we found these to be negatively correlated (i.e. β_i and γ_i in the baseline model
 312 were positively correlated; Fig. 2). This means that species found at lower elevations have
 313 generally wider distributions than those at higher elevations. The same relationship was
 314 found when using instead elevation or mean temperature as explanatory variables (Supple-
 315 mentary Fig. XX) as well as when using ordinal data (Supplementary Fig. XX); however,
 316 the pattern was not present along the second axis of variation of the environmental variables
 317 (Supplementary Fig. XX). The comparison between the other parameter estimates revealed
 318 additional, somewhat more expected, relationships. In particular, we found the amplitude
 319 of distributions to be positively and negatively correlated with their mean and the variance,
 320 respectively (i.e. α_i is positively correlated with β_i and γ_i ; Supplementary Fig. XX). This
 321 implies that, at higher elevations, species' distributions generally have lower amplitudes.

322 Maintaining the symmetry of species' distributions, we then allowed the kurtosis—or shape
 323 of the tails—of these to vary in different ways. To do so, we changed the response curve of our
 324 Bayesian model to follow a generalized error distribution (Eq. 5). A comparison of the WAIC
 325 values showed this non-linear regression to outperform the baseline model (Supplementary
 326 Fig. XX). Studying the resulting posterior distributions, we found the average kurtosis of the
 327 distributions to be slightly greater than zero, which corresponds to distributions with longer
 328 tails than the Gaussian case (Fig. 3). However, the parameter controlling for the kurtosis
 329 ν_i displayed a lot of variation across species (Supplementary Fig. XX), which might indicate
 330 that the shape of the tails is species-specific.

331 Using Eq. (6), we next studied the skewness of species' distributions. Based on the esti-
 332 mates for the WAIC values, this model outperformed the first two (Supplementary Fig. XX),
 333 which sheds light on the naturally skewed nature of species' distributions. Perhaps most im-
 334 portantly, studying the mean value of the skewness across species, we found this to be
 335 consistently below zero (Fig. 3). This indicates that species' distributions generally present
 336 steeper declines towards higher elevations (i.e. $\hat{\lambda} < 0$; Fig. 1). The same was true when using
 337 a model that allowed for both fat-tailed and skewed response curves (Eq. 7). This model out-
 338 performed the rest, presenting Akaike weights close to 1 (Supplementary Fig. XX), suggesting

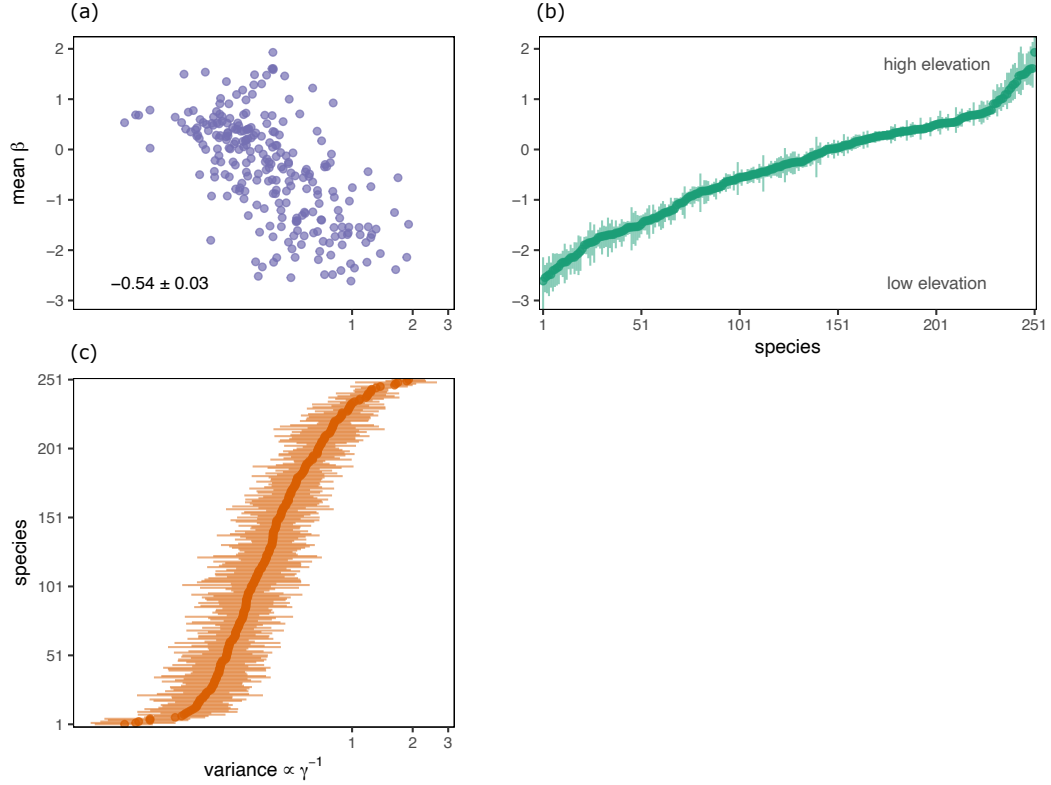


Figure 2: Relationship between mean and variance of species' distributions. Posterior distributions for parameters β_i and γ_i from Eq. (1) across species, and the relationship between them. Panel (a) describes the relationship between range size and elevation. Every dot represents the relationship between the mean values for the β_i and γ_i estimates of the different species. The value in the top-right corner of the plot displays the Pearson's correlation between these parameters calculated across samples of the posterior distributions. Panel (b) describes the β_i posterior distribution estimated for all species. Panel (c) describes the γ_i posterior distribution estimated for all species. In (b) and (c), the points represent the mean of the posterior distributions, and the corresponding lines characterize the 89% confidence intervals.

that both the kurtosis and skewness are useful properties to describe empirical distributions (Fig. 3).

The model characterizing fat-tailed and skewed distributions allowed us to study the posterior distributions for the parameters describing the mean, variance, amplitude, kurtosis and skewness of species realized niches altogether. We observed that different type of species seem to present characteristically different distributions (Fig. XX). For example, focussing on the aforementioned negative correlation between the mean and variance of species' distributions, we found neophyte and archaeophyte species to escape this apparent macroecological constraint (Fig. XX). Similarly, neophytes' distributions also appear to showcase steeper declines towards lower elevations, which is in contrast to the general pattern found across

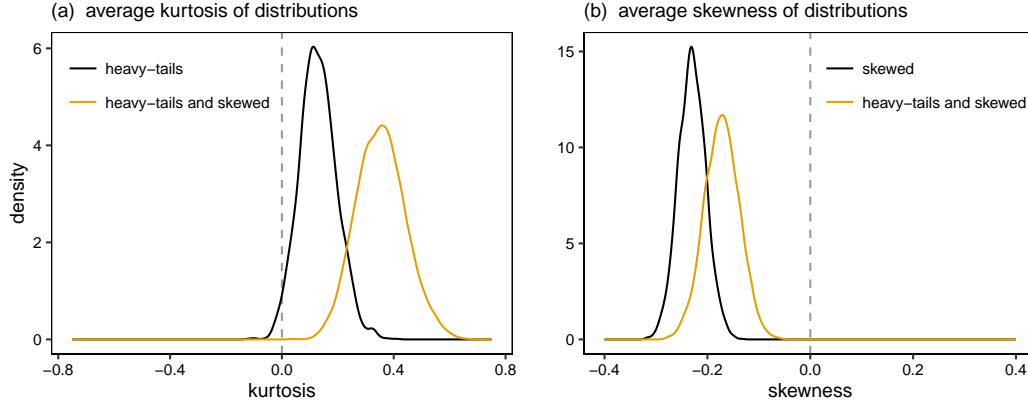


Figure 3: Average kurtosis and skewness of species' distributions. Calculated using the posterior distributions of parameters $\hat{\nu}$ and $\hat{\lambda}$ from the models (see Supplementary Table XX), the two panels describe the average (a) kurtosis and (b) skewness of distributions, respectively. Panel (a) displays the results obtained by using a response curve that follows a generalized error distribution. Panel (b) displays the results obtained by using a response curve that follows a skewed normal distribution. In both case, the red dotted line indicates the conditions by which species are normally distributed along the environmental axis.

species. Notice the nature of these results does not depend on the presence or absence of a species at the edge of the sampling area, as the same model produced comparable results when using simulated and bootstrapped data (Supplementary Information and Supplementary Fig. XX). The comparison across other types of species also provided us with additional interesting mappings. For example,... Likewise, other pairwise comparisons across parameters also revealed additional macroecological constraints (Supplementary Fig. XX).

Discussion

In this work, we used non-linear response curves to model the distribution of species across an environmental gradient. First, we used a baseline model that considered these as bell-shaped, and we studied the relationship between the basic parameters characterizing them. We found both the amplitude and variance of distributions to be negatively correlated with elevation. Considering more complex response curves, we then found species' distributions to present non-normal tails and skewed shapes. That is, species' distributions were generally characterized by fat tails and steeper declines towards higher elevations. That said, the nature of these distributions was not homogeneous across species, as some species presented singularly different properties. This is the case of neophyte, which present distributions ...

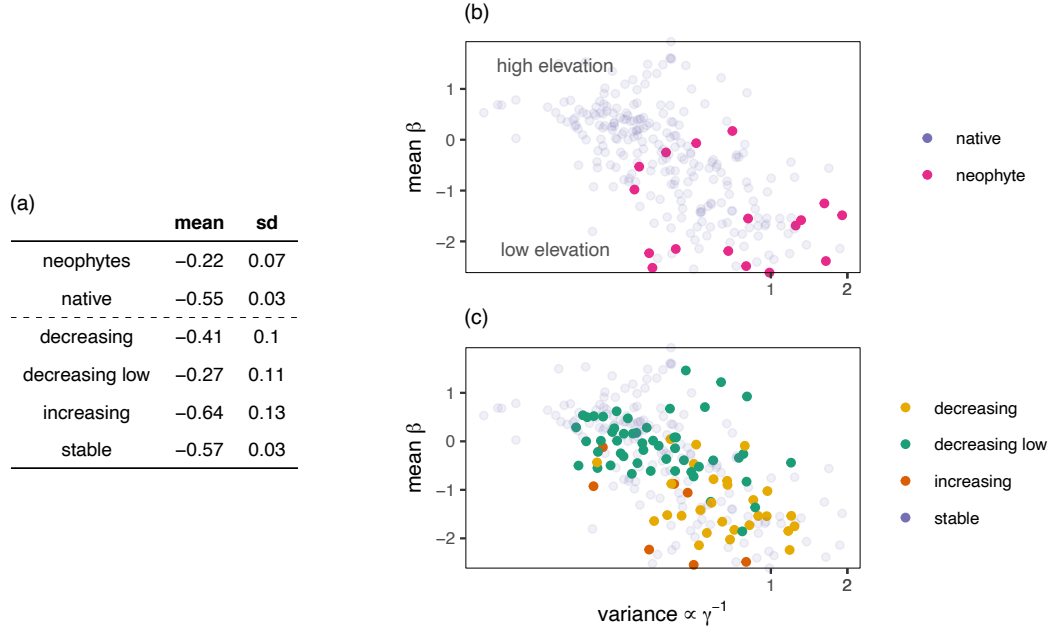


Figure 4: Universality of the relationship between mean and variance of species’ distributions. Comparison between how different types of species are mapped in Fig. 2a. Panel (a) describes the correlation coefficient between β_i and γ_i for each type of species. Panel (b) shows the differences between neophytes and native species in the way these are distributed along the environmental gradient. Panel (c) shows the same differences for species that have decreased, decreased in low elevations, increase and remain stable over the last decades (see Supplementary Table 1 for further details).

The statistical framework used here allowed us to parsimoniously compare the shape of the species’ realized niches along an altitude gradient, testing for the existence of several interesting macroecological patterns. We first studied the relationship between the mean and variance of species’ distributions. The Rapoport’s rule predicts wider ranges of species at higher latitudes and altitudes (Stevens, 1992); and therefore, one might expect a positive correlation between these properties. A common explanation for the Rapoport’s rule is that climatic variability selects for species with greater climatic tolerances. However, 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 (Ribas & Schoereder, 2006; Bhattarai & Vetaas, 2006; McCain & Knight, 2013). Our results seem to contradict the predictions of the Rapoport’s rule, as we observed a negative correlation between species’ range and elevation.

Similarly, we study whether or not species’ distributions show steeper declines towards stressful conditions, testing the so-called abiotic stress limitation hypothesis (?). Normand

et al. (2009) tested this for vegetation data using Huisman *et al.*'s statistical models for several independent species, finding no clear support for such a hypothesis. Our results are able to shed light on this geographical pattern as well as to highlight the degree to which different species will showcase different levels of decline towards stressful conditions. Specifically, we are able to link plant physiological traits to the skewness of their distributions.

Structure of the discussion section:

1. Summary of results.
2. Proper test of Rapopor's hypothesis. Different species follow different biogeographical patterns.
3. Proper test of skewed towards high altitude. Is species' physiology informative to explain the pattern?
4. What is the true shape of species' distributions? These display heavy-tail and skewed distributions.
5. Future directions. Missing bimodal curves. Using this information to understand where jSDMs estimate interactions between species. Further test of the ability of traits to predict those parameter estimates.

First, we test for the Rapopor'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 (Ribas & Schoereder, 2006; Bhattarai & Vetaas, 2006; McCain & Knight, 2013). Our results not only allow us to properly test the existence of this geographical pattern, but they also showcase variation in how different types of species, such as native or neophytes, might respond to an environmental gradient.

Second, we study whether or not species' distributions show steeper declines towards stressful conditions, testing the so-called abiotic stress limitation hypothesis (?). Normand *et al.* (2009) tested this for vegetation data using Huisman *et al.*'s statistical models for several independent species, finding no clear support for such a hypothesis. Our results are able to shed light on this geographical pattern as well as to highlight the degree to which different

407 species will showcase different levels of decline towards stressful conditions. Specifically, we
408 are able to link plant physiological traits to the skewness of their distributions.

409 There is a lot of variation that remains unexplained regarding the shape of species' distri-
410 bution, and Fig... gives us glimpse of it.

References

- Ashour, S. K. & Abdel-hameed, M. A. (2010). Approximate skew normal distribution. *Journal of Advanced Research*, 1, 341–350.
- Austin, M. P. (1987). Models for the analysis of species’ response to environmental gradients. *Vegetatio*, 69, 35–45.
- Austin, M. P. (2002). Spatial prediction of species distribution: An interface between ecological theory and statistical modelling. *Ecological Modelling*, 157, 101–118.
- Bhattarai, 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.
- 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.
- Evans, M. E. K., Merow, C., Record, S., McMahon, S. M. & Enquist, B. J. (2016). Towards Process-based Range Modeling of Many Species. *Trends in Ecology & Evolution*, 31, 860–871.
- 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.
- Guisan, A. & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–186.
- 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.
- Jamil, T. & ter Braak, C. J. F. (2013). Generalized linear mixed models can detect unimodal species-environment relationships. *PeerJ*, 1, e95.

- 437 Kaufman, D. M. (1995). Diversity of New World Mammals: Universality of the Latitudinal
438 Gradients of Species and Bauplans. *Journal of Mammalogy*, 76, 322–334.
- 439 Krebs, C. J. (1972). *Ecology: The Experimental Analysis of Distribution and Abundance/by*
440 *Charles J. Krebs*. 4th edn. Harper & Row, New York.
- 441 Landolt, E., Bäumler, B., Ehrhardt, A., Hegg, O., Klötzli, F., Lämmler, W., Nobis, M.,
442 Rudmann-Maurer, K., Schweingruber, F. H., Theurillat, J.-P., Urmi, E., Vust, M. &
443 Wohlgemuth, T. (2010). *Flora indicativa: Ökologische Zeigerwerte und biologische Kennze-*
444 *ichen zur Flora der Schweiz und der Alpen*. Haupt, Bern. ISBN 978-3-258-07461-0.
- 445 Malmgren, R. D., Stouffer, D. B., Motter, A. E. & Amaral, L. A. N. (2008). A Poissonian
446 explanation for heavy tails in e-mail communication. *Proceedings of the National Academy*
447 *of Sciences*, 105, 18153–18158.
- 448 McCain, C. M. & Knight, K. B. (2013). Elevational Rapoport’s rule is not pervasive on
449 mountains. *Global Ecology and Biogeography*, 22, 750–759.
- 450 Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics*, 32,
451 685–694.
- 452 Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., Araújo,
453 M. B., Dallas, T., Dunson, D., Elith, J., Foster, S. D., Fox, R., Franklin, J., Godsoe, W.,
454 Guisan, A., O’Hara, B., Hill, N. A., Holt, R. D., Hui, F. K. C., Husby, M., Kålås, J. A.,
455 Lehtikainen, A., Luoto, M., Mod, H. K., Newell, G., Renner, I., Roslin, T., Soininen, J.,
456 Thuiller, W., Vanhatalo, J., Warton, D., White, M., Zimmermann, N. E., Gravel, D. &
457 Ovaskainen, O. (2019). A comprehensive evaluation of predictive performance of 33 species
458 distribution models at species and community levels. *Ecological Monographs*, 89, e01370.
- 459 Normand, S., Treier, U. A., Randin, C., Vittoz, P., Guisan, A. & Svenning, J.-C. (2009).
460 Importance of abiotic stress as a range-limit determinant for European plants: Insights
461 from species responses to climatic gradients. *Global Ecology and Biogeography*, 18, 437–449.
- 462 Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F. G., Duan, L., Dunson, D., Roslin,
463 T. & Abrego, N. (2017). How to make more out of community data? A conceptual
464 framework and its implementation as models and software. *Ecology Letters*, 20, 561–576.

- 465 Petrovskii, S., Morozov, A., Taylor, A. E. P. D. & DeAngelis, E. D. L. (2009). Dispersal in
466 a Statistically Structured Population: Fat Tails Revisited. *The American Naturalist*, 173,
467 278–289.
- 468 Ribas, C. R. & Schoereder, J. H. (2006). Is the Rapoport effect widespread? Null models
469 revisited. *Global Ecology and Biogeography*, 15, 614–624.
- 470 Rohde, K. (1992). Latitudinal Gradients in Species Diversity: The Search for the Primary
471 Cause. *Oikos*, 65, 514–527.
- 472 Scherrer, D. & Guisan, A. (2019). Ecological indicator values reveal missing predictors of
473 species distributions. *Scientific Reports*, 9, 1–8.
- 474 Stan Development Team (2021). RStan: The R interface to Stan.
- 475 Stan Development Team (2021). Stan Modeling Language Users Guide and Reference Man-
476 ual.
- 477 Stevens, G. C. (1992). The Elevational Gradient in Altitudinal Range: An Extension of
478 Rapoport’s Latitudinal Rule to Altitude. *The American Naturalist*, 140, 893–911.
- 479 Sunday, J. M., Bates, A. E. & Dulvy, N. K. (2011). Global analysis of thermal tolerance
480 and latitude in ectotherms. *Proceedings of the Royal Society B: Biological Sciences*, 278,
481 1823–1830.
- 482 ter Braak, C. J. F. & Looman, C. W. N. (1986). Weighted averaging, logistic regression and
483 the Gaussian response model. *Vegetatio*, 65, 3–11.
- 484 Wong, F. & Collins, J. J. (2020). Evidence that coronavirus superspreading is fat-tailed.
485 *Proceedings of the National Academy of Sciences*, 117, 29416–29418.