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

Bernat Bramon Mora[1,*] and Jake M. Alexander[1]

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

Abstract

Species' distribution models have emerged as one of the most influential methodological advances in ecology and biogeography of the last decades. Useful to understand how populations of species will change along environmental gradients, they have become ecologists' compass to predict the effects of global climate change. That said, uncovering the mechanisms shaping species' realized niches has been one of the main driving forces behind the development of these models. That is, recent efforts have been often focused on understanding which biotic and abiotic factors are good predictors of species' niches—with an increasing effort placed in improving the predictive power of the statistical models. However, we still lack a general understanding of the shape of species' distributions, and much less is known about how these distributions compare to each other across gradients. Here, we use a set of Bayesian non-linear models to uncover the shape of species' realized niches. These models account for all prior knowledge we have regarding their shape, including expert knowledge on species' environmental preferences and physiology. With this approach, we are able to uncover the true shape of empirical species' distributions. Moreover, they allow us to tackle long-standing hypothesis regarding general biogeographical patterns. In particular, we found conclusive evidence of the relationship between several properties of distributions, including the link between species' range size and elevation and their skewness along gradients. Finally, we are able to shed light on the extent to which some aspects of the shape of observed realized niches—such as kurtosis and skewness of the distributions—could be intrinsic properties of species. Overall, our approach offers a useful statistical framework to understand

the shape of species' distributions, and our results provide an unprecedented perspective of the way systems of many species are distributed along environmental gradients.

Introduction

One of the central goals of ecology is to understand the ways species are distributed across space and time (ref). Over the last two decades, ecologists have developed multiple distribution models to try to untangle the factors that play a role in defining such distributions (Guisan & Zimmermann, 2000). These models estimate species' realized niches using several covariates, including environmental variables (Guisan & Thuiller, 2005), species ecological traits' (Pollock *et al.*, 2012) and phylogenetic relations (Ives & Helmus, 2011). More recently, some of the focus have shifted towards approaches that estimate and account for biotic factors, such as competitive or facilitative relationships between species (Ovaskainen *et al.*, 2017). The idea is that by untangling the ways in which such biotic and abiotic factors shape species' distributions, we can gain a 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 recently questioned (Harris, 2016; Thurman *et al.*, 2019; Poggiato *et al.*, 2021). 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 dispersal ability in setting species' range limits (Wisz *et al.*, 2013; Pollock *et al.*, 2014; 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 (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 (Kaufman, 1995). 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 (Krebs, 1972) 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, 2002). In practice, distribution frameworks often use logistic regressions with a linear relationship between covariates (but see **?** and **?**). This is useful because it simplifies the optimization process, but it comes with some statistical shortcomings. First and foremost, such response curve and the linear relationship between covariates often comes with a set of implicit mathematical constrains that might not be biologically 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). Other factors might then condition species distributions to showcase heavy-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 constrains 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; Ovaskainen *et al.* 2017). 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 biogeoraphy. In particular, we develop a Bayesian hierarchical model that accounts for all prior information that we have regarding the distribution of alpine plant species along an elevation gradient in the Swiss Alps, including expert knowledge on species environmental indicator values, range sizes, and plant physiology. We start by considering species' response curves as Gaussian distributed, and then we adapt our model to allow for skewed and long-tailed distributions. Using this 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 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 (ref). 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. Overall, we use models that are solely constrained by the empirical information that we truly have regarding our system, relaxing as much as possible the structural constrains of the statistical framework. Using these models, we are able 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 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 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 climatic 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' physiological traits 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 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 types (e.g. identifying those species that are neophytes) 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 characterize 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 specie in the site—we do not necessarily have an *a priori* expectation of how exactly these factors will influence the shape of species' distributions. Therefore, for this baseline model, if all that we are willing to assume about species' realized niches is that these have finite variance, the most conservative assumption and the safest bet—i.e. the one with the largest entropy—is that they follow a Gaussian distribution (Fig 1a). That is, given the presence/absence or abundance y_{ij} of any species i in any given site j, and an environmental variable x_j, we can define species' responses to the environment as

$$\begin{aligned} y_{ij} &\sim F\left(p_{ij}\right) \\ \log\left(p_{ij}\right) &= -\alpha_i - \gamma_i \left(x_j - \beta_i\right)^2, \end{aligned} \tag{1}$$

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

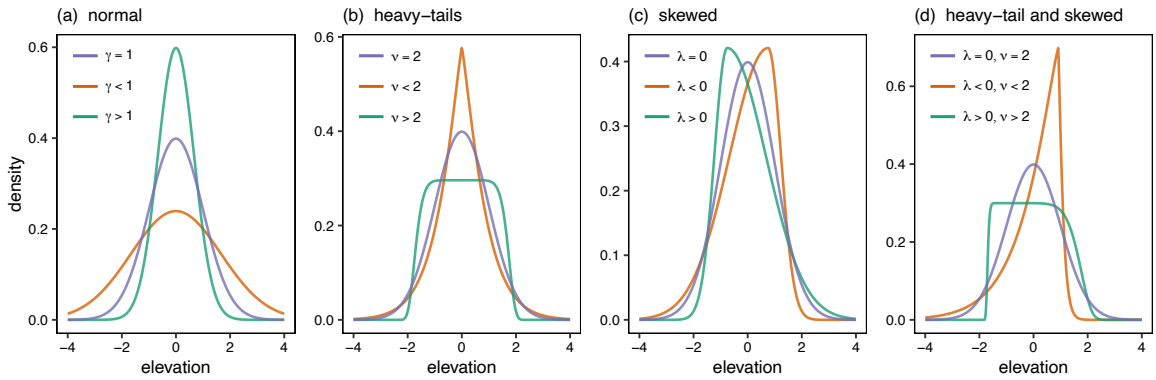

Figure 1: Different response curves. Panel (a) shows the probability density function characterized by Eq. (1) for different values of γ, when $\mu = 0$. Panel (b) shows the probability density function characterized by Eq. (5) for different values of ν, when $\gamma = 1$ and $\mu = 0$. Panel (c) shows 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.

Model priors

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

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

where parameters γ_i and β_i are expressed as multivariate normal distributions—i.e. Gaussian processes—such that Σ^{β} and Σ^{γ} are variance-covariance matrices describing species' similarity in terms of their average climatic suitability and range of variation along the different environmental gradients, respectively. We define these variance-covariance matrices as follows:

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

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

$$\Sigma_{ij} = \eta \exp\left(-\sum_k \rho_k {D^k_{ij}}^2\right) + \delta_{ij}\sigma, \tag{4}$$

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 Developent 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 binary data. When using ordinal data, we limit our study to a random subset of a 100 species of those with at least 20 occurrences to facilitate the sampling of the models.

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

Heavy-tail response curve

Distributions with heavy-tails are very common across fields, as they can capture processes such as seasonality (e.g. in communications patterns; Malmgren *et al.* 2008) or some stochastic events (e.g. in the spread of infectious diseases; Wong & Collins 2020). Indeed, heavy-tail distributions are pervasive in ecology; for example, species' dispersal patterns have been shown to have heavy-tails due to natural variability among individuals (Petrovskii *et al.*, 2009). Therefore, one might expect these properties to also emerge in alpine communities, where seasonality and dispersal patterns are crucial factors determining species' distributions. To accommodate this feature into our baseline model, we could consider a response curve that follows a generalized error distribution, since the Gaussian shape is a special case

of it and contains parameter that regulates the level of kurtosis (Fig 1b). In particular, we can adapt Eq. (1) to present this non-linear form as follows:

$$\log\left(p_{ij}\right) = -\alpha_i - \gamma_i' \left|x_j - \beta_i\right|^{\nu_i}, \tag{5}$$

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 $\log\left(\nu_i - 1\right) \sim \text{Normal}\left(\hat{\nu}, \sigma_\nu\right)$, where $\hat{\nu} \sim \text{Normal}\left(0, 1\right)$ and $\sigma_\nu \sim \text{Exponential}\left(2\right)$. 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 heavy-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 heavy 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

When species experience abiotic or biotic pressures that increase or decrease along an environmental gradient, one might expect their distributions to be skewed in one direction. Likewise, this same skewed shape can also emerge as a result of species' asymmetric environmental tolerance. One way to accommodate this feature to our models is by considering a skewed normal distribution (Supplementary Figure XX). As for the two cases 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

$$p_{ij} = \hat{p}_{ij}\left[1 + \text{erf}\left(\lambda_i\left(x_j - \beta_i'\right)\sqrt{\frac{\gamma_i'}{2}}\right)\right]$$
$$\log\left(\hat{p}_{ij}\right) = -\alpha_i' - \gamma_i'\left(x_j - \beta_i'\right)^2, \quad (6)$$

where $\alpha_i' = q_1\left(\alpha_i, \gamma_i', \lambda_i\right)$, $\beta_i' = q_2\left(\gamma_i', \lambda_i\right)$, $\gamma_i' = q_3\left(\gamma_i, \lambda_i\right)$, and λ_i is a parameter that 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 Ashour & Abdel-hameed 2010). We choose an adaptive prior such that $\lambda_i \sim \text{Normal}\left(\hat{\lambda}, \sigma_\lambda\right)$, where $\hat{\lambda} \sim \text{Normal}\left(0, 1\right)$ and $\sigma_\lambda \sim \text{Exponential}\left(1\right)$. This model can be re-parametrized following q_1, q_2 and q_3, and 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).

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 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.

Heavy-tail and skewed response curve

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 (also see XX). 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 **1**d). The skewed generalized error distribution can be considered by modifying

the species' response curve in Eq. (1) as

$$\log\left(p_{ij}\right) = -\alpha_i - \left(\frac{\gamma_i' \left|x_j - \beta_i'\right|}{1 + \lambda_i \operatorname{sgn}\left(x_j - \beta_i'\right)}\right)^{\nu_i}, \tag{7}$$

where $\gamma_i' = f_1\left(\gamma_i, \nu_i, \lambda_i\right)$, $\beta_i' = f_2\left(\gamma_i, \beta_i, \nu_i, \lambda_i\right)$, ν_i and λ_i are parameters that describe the kurtosis and skewness of the distribution, respectively. The function $\operatorname{sgn}(x)$ characterizes the sign function. We define ν_i and its prior as in Eq. 5. However, λ_i is now defined as $\lambda_i \in (-1, 1)$, with an adaptive prior such that $\operatorname{logit}\left(\frac{\lambda_i+1}{2}\right) \sim \text{Normal}\left(\hat{\lambda}, \sigma_\lambda\right)$, where $\hat{\lambda} \sim \text{Normal}(0, 1)$ and $\sigma_\lambda \sim \text{Exponential}(1)$. 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.

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 and absence of species across sites as the 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 Rapopor'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. **2**). 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 (Fig. **3**). 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).

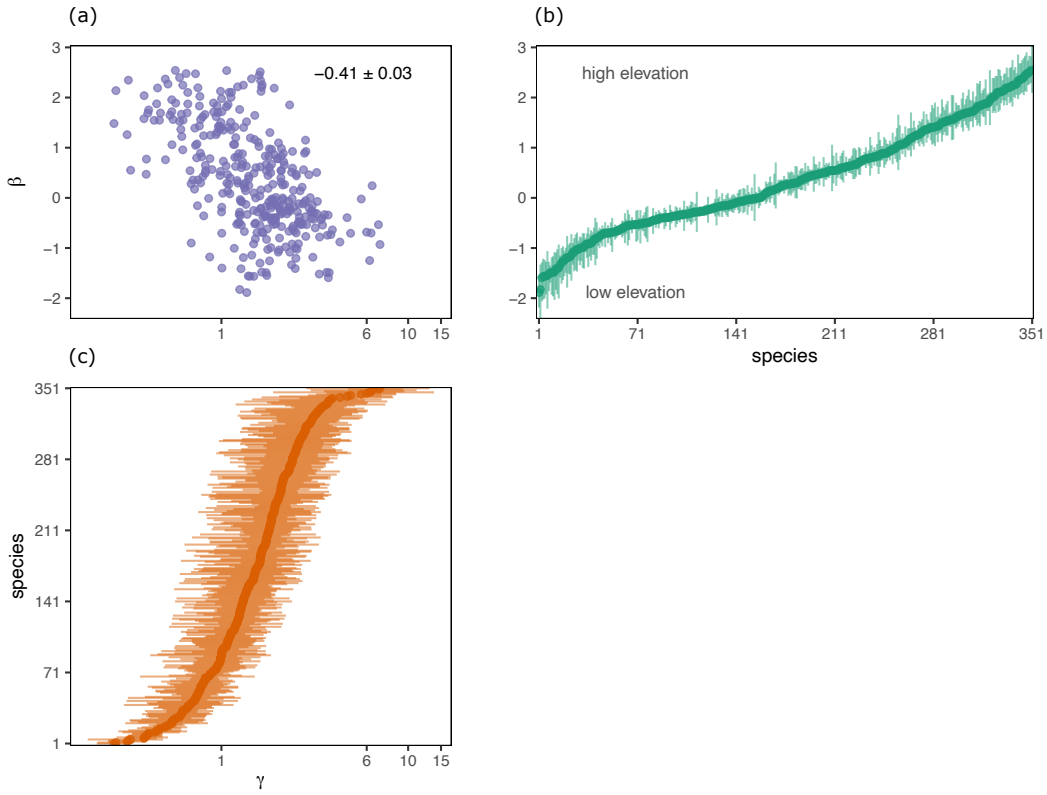

Figure 2: Relationship between mean and variance of species' distributions. Posterior distributions for parameters β_i and γ_i from Eq. (1) across species, and the relationship between them. Panel (a) describes the relationship between range size and elevation. Every dot represents the relationship between 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.

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 a generalized error distribution (Eq. 5). A comparison of the WAIC values showed this non-linear regression to outperform the baseline model (Supplementary Fig. XX)[†]. Studying the resulting posterior distributions, we found the parameter controlling for the kurtosis to be centred around $\nu_i \sim 2$, which corresponds to a distribution with a Gaussian shape (Fig. 4). However, such parameter estimates displayed a lot of variation across species, which might indicate that the shape of the tails of the distribution is species-

[†] I am still waiting on the results for this (currently running in the cluster), and this is just my prior expectation based on what I've seen in some of the other models I've been working with.

(a)

	mean	sd
neophytes	−0.4	0.1
native	−0.4	0.03
decreasing	−0.31	0.11
decreasing low	−0.19	0.07
increasing	−0.48	0.14
stable	−0.43	0.03

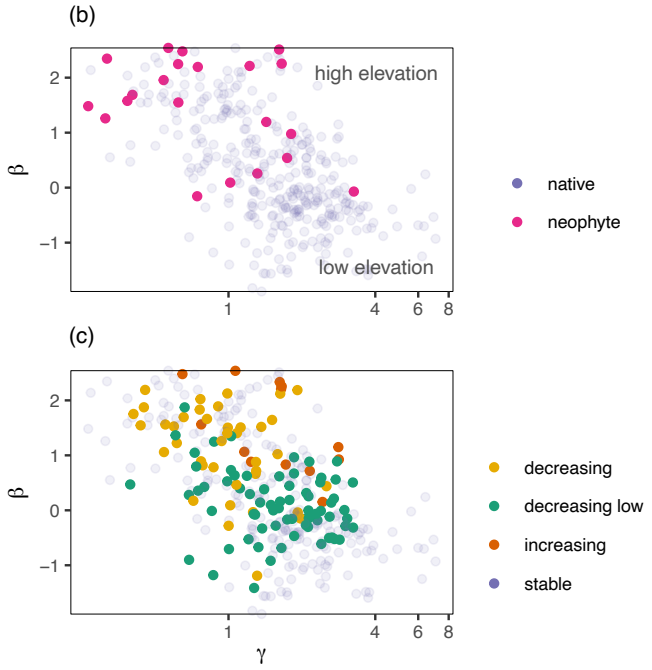

Figure 3: Universality of the relationship between mean and variance of species' distributions. Comparison between how different types of species are mapped in Fig. 2a. Panel (a) describes the correlation coefficient between β_i and γ_i for each type of species. Panel (b) shows the differences between 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). I might move this to the Supplementary Information.

specific (Supplementary Fig. XX).

Using Eq. (6), we next studied the skewness of species distributions. Based on the estimates for the WAIC values, this model outperformed the rest (Supplementary Fig. XX), which sheds light on the naturally skewed nature of species' distributions. Perhaps most importantly, studying the mean value of the skewness across species (i.e. parameter $\hat{\lambda}$), we found that species' distributions present steeper declines towards lower elevations (Fig. 4). Moreover, species' physiology seems to strongly shape this parameter, which suggest that distribution skewness is an intrinsic property of species (New Fig. XX)†.

Finally, we used a model that allowed for heavy-tails and skewed response curves (Eq. 7) to understand the relationship between the kurtosis and skewness of distributions. This model outperformed the rest, presenting Akaike weights close to 1 when comparing models (Supplementary Fig. XX). Comparing the posterior distribution for parameters ν_i and λ_i, we found this to be uncorrelated†. However, different types of species seem to present

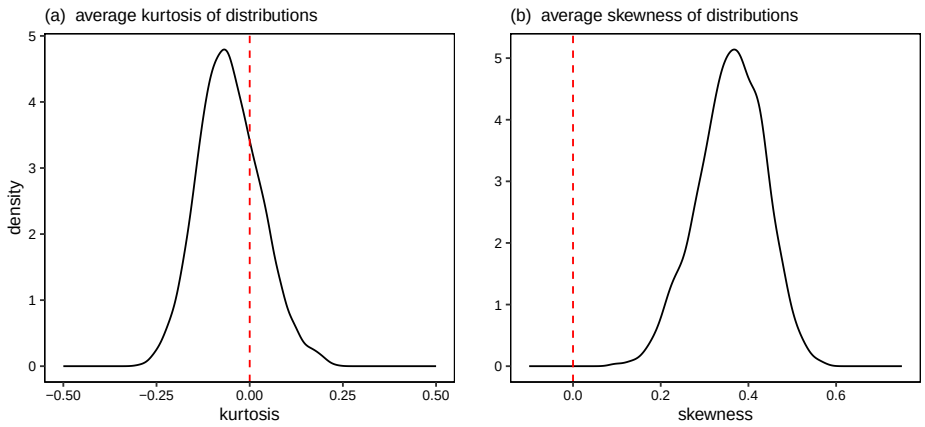

Figure 4: Average kurtosis and skewness of species' distributions. Calculated using the posterior distributions of parameters $\hat{\nu}$ and $\hat{\lambda}$ from the models (see Supplementary Table XX), the two panels describe the average (a) kurtosis and (b) skewness of distributions, respectively. 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.

characteristically different distributions based on their kurtosis and skewness. Importantly, this model identified similar relationships between parameters as the baseline model (Supplementary Fig. XX), and measured similar average levels of kurtosis and skewness than the the other two models (Fig. 4)[†].

Discussion

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.

References

Ashour, S. K. & Abdel-hameed, M. A. (2010). Approximate skew normal distribution. *Journal of Advanced Research*, 1, 341–350.

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.

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.

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.

Guisan, A. & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, 8, 993–1009.

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.

Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*, vol. 846. John Wiley & Sons.

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., Lämmler, W., Nobis, M., Rudmann-Maurer, K., Schweingruber, F. H., Theurillat, J.-P., Urmi, E., Vust, M. & Wohlgemuth, T. (2010). *Flora indicativa: Okologische 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.

Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics*, 32, 685–694.

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

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

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

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

Poggiato, G., Münkemüller, T., Bystrova, D., Arbel, J., Clark, J. S. & Thuiller, W. (2021). On the Interpretations of Joint Modeling in Community Ecology. *Trends in Ecology & Evolution*.

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

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

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

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

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

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

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

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

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

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

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

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

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