Received: 19 May 2023 | Accepted: 26 February 2024

DOI: 10.1111/2041-210X.14312

Methods in Ecology and Evolution

RESEARCH ARTICLE

Bayesian hierarchical modelling of size spectra

Jeff S. Wesner[1] | Justin P. F. Pomeranz[2] | James R. Junker[3,4] | Vojsava Gjoni[1]

[1]Department of Biology, University of South Dakota, Vermillion, South Dakota, USA

[2]Department of Physical and Environmental Sciences, Colorado Mesa University, Grand Junction, Colorado, USA

[3]Great Lakes Research Center, Michigan Technological University, Houghton, Michigan, USA

[4]Department of Biological Sciences, Advanced Environmental Research Institute, University of North Texas, Denton, Texas, USA

Correspondence
Jeff S. Wesner
Email: jeff.wesner@usd.edu

Funding information
National Science Foundation, Grant/Award Number: 2106067 and 2106068

Handling Editor: Fernando González Taboada

Abstract

1. A fundamental pattern in ecology is that smaller organisms are more abundant than larger organisms. This pattern is known as the individual size distribution (ISD), which is the frequency distribution of all individual body sizes in an ecosystem.
2. The ISD is described by a power law and a major goal of size spectra analyses is to estimate the exponent of the power law, λ. However, while numerous methods have been developed to do this, they have focused almost exclusively on estimating λ from single samples.
3. Here, we develop an extension of the truncated Pareto distribution within the probabilistic modelling language Stan. We use it to estimate multiple λs simultaneously in a hierarchical modelling approach.
4. The most important result is the ability to examine hypotheses related to size spectra, including the assessment of fixed and random effects, within a single Bayesian generalized mixed model. While the example here uses size spectra, the technique can also be generalized to any data that follow a power law distribution.

KEYWORDS

Bayesian, body size spectra, hierarchical, Pareto, power law, Stan

1 | INTRODUCTION

In any ecosystem, large individuals are typically more rare than small individuals. This fundamental feature of ecosystems leads to a remarkably common pattern in which relative abundance declines with individual body size, generating the individual size distribution (ISD), also called the community size spectrum (Platt & Denman, 1977; Sprules et al., 1983; White et al., 2008). Understanding how body sizes are distributed has been a focus in ecology for over half a century (Kerr, 1974; Peters & Wassenberg, 1983; Sheldon & Parsons, 1967), in part because body size distributions reflect fundamental measures of ecosystem structure and function, such as trophic transfer efficiency (Kerr & Dickie, 2001; Perkins et al., 2019; White et al., 2007). Individual size distributions are also predicted as a result of physiological limits associated with body size, thereby emerging from predictions of metabolic theory and energetic equivalence (Brown et al., 2004).

More formally, the ISD can be modelled as a probability density function with a single free parameter λ:

$$f(x) = Cx^{\lambda}, \quad x_{min} \le x \le x_{max} \tag{1}$$

where x is the body size (e.g. mass or volume) of an individual, x_{min} is the smallest possible individual and x_{max} is the largest possible individual. C is a constant equal to:

$$C = \begin{cases} \dfrac{\lambda+1}{x_{max}^{\lambda+1} - x_{min}^{\lambda+1}}, \lambda \neq -1 \\ \dfrac{1}{\log x_{max} - \log x_{min}}, \lambda = -1 \end{cases} \quad (2)$$

This model is also known as the bounded power law or truncated Pareto distribution. The term 'bounded' or 'truncated' refers to the minimum x_{min} and maximum x_{max} possible body sizes (Edwards et al., 2017; White et al., 2008).

A compelling feature of size spectra is that λ may vary little across ecosystems as a result of physiological constraints that lead to size-abundance patterns more broadly. Metabolic scaling theory predicts $\lambda + 1 = \frac{\log_{10}\alpha}{\log_{10}\beta} - 3/4$, where α is trophic transfer efficiency in the food web and β is the mean predator–prey mass ratio (Reuman et al., 2008). The value of $-3/4$ is the scaling coefficient of metabolic rate and mass (0.75) (Brown et al., 2004), and as a result, values of λ have been used to estimate metabolic scaling across ecosystems (Perkins et al., 2018, 2019; Reuman et al., 2008). Values of ~−2 represent a reasonable first guess of expected ISD exponents, with values of ranging from −1.2 to −2 appearing in the literature (Andersen & Beyer, 2006; Blanchard et al., 2009; Pomeranz et al., 2022).

Whether λ represents a fixed or variable value is debated, but it varies among samples and ecosystems (Blanchard et al., 2009; Perkins et al., 2018; Pomeranz et al., 2022). It is often described by its connection with the steepness of log–log plots of size spectra, with more negative values (i.e. 'steeper') indicating lower abundance of large relative to small individuals, and vice versa. These patterns of size frequency are an emergent property of demographic processes (e.g. age-dependent mortality), ecological interactions (e.g. size-structured predation, trophic transfer efficiency) and physiological constraints (e.g. size-dependent metabolic rates; Andersen & Beyer, 2006; Muller-Landau et al., 2006; White et al., 2008). As a result, variation in λ across ecosystems or across time can indicate fundamental shifts in community structure or ecosystem functioning. For example, overfishing in marine communities has been detected using the size spectrum in which λ was steeper than expected, indicating fewer large fish than expected (Jennings & Blanchard, 2004). Shifts in λ have also been used to document responses to acid mine drainage in streams (Pomeranz et al., 2019), land use (Martínez et al., 2016), resource subsidies (Perkins et al., 2018) and temperature (O'Gorman et al., 2017).

Given the ecological information it conveys, the data required to estimate size spectra—a vector of individual body sizes—are deceptively simple. As long as the body sizes are collected systematically and without bias towards certain sizes, there is no need to know any more ecological information about the data points (e.g. trophic position, age, abundance). However, the statistical models used to estimate λ are diverse. Edwards et al. (2017) documented eight methods. Six involved binning, in which the body sizes are grouped into size bins (e.g. 2–50 mg, 50–150 mg, etc.) and then counted, generating values for abundance within each size bin. Binning and log transformation allows λ to be estimated using simple linear regression. Unfortunately, the binning process also removes most of the variation in the data, collapsing information from 1000s of individuals into just six or so bins. Doing so can lead to inaccurate values of λ, sometimes drastically so (Goldstein et al., 2004; Pomeranz et al., 2024; White et al., 2008).

An improved alternative to binning and linear regression is to fit the body size data to a power law probability distribution (Edwards et al., 2017, 2020; White et al., 2008). This method uses all raw data observations directly to estimate λ using the maximum likelihood estimation method (Edwards et al., 2017). In addition to estimating size spectra of single samples, ecologists have used this method to examine how λ varies across environmental gradients (Perkins et al., 2019; Pomeranz et al., 2022). However, these analyses typically proceed in two steps. First, λ is estimated individually from each collection (e.g. each site or year, etc.). Second, the estimates are used as response variables in a linear model to examine how they relate to corresponding predictor variables (Edwards et al., 2020). We refer to this as the 'two-step' approach. A downside to the two-step approach is that it treats body sizes (and subsequent λs) as independent samples, even if they come from the same site or time. It also removes information on sample size (number of individuals) used to derive λ. As a result, the approach not only separates the data generation model from the predictor variables but also it is unable to take advantage of partial pooling, in which group-level estimates exhibit shrinkage towards each other (Gelman et al., 2012).

Here, we develop a Bayesian modelling framework that uses the truncated Pareto distribution to estimate λ in response to both fixed and random predictor variables. The primary benefit of this approach is that it combines the data generation process and the linear (or non-linear) model into a single generalized linear (or non-linear) mixed model. The model extends the maximum likelihood approach developed by Edwards et al. (2020) to allow for a flexible hierarchical structure, including partial pooling, within the modelling language Stan (Stan Development Team, 2022).

2 | METHODS

2.1 | Translating to Stan

Stan is a probabilistic modelling language that estimates Bayesian posteriors using Hamiltonian Monte Carlo (Stan Development Team, 2022). It does not contain the truncated Pareto described in Equation 2, so we added it as a user-defined log probability density function (lpdf) in `rstan` (Annis et al., 2017; Stan Development

Team, 2022) (Supporting Information S1). The lpdf was slightly modified as described in S.1.4 of Edwards et al. (2020) to contain a term for the count of each body size. In data sets with individual body sizes, counts will be a simple constant with a value of 1. However, if a sample of body sizes is $x=\{1.2, 1.2, 1.5, 2.8\}$, then these can be re-formatted to include a vector for all unique x values {1.2, 1.5, 2.8} and a vector for their counts = {2,1,1}, where counts can also be non-integers. In large data sets (e.g. >500 individuals or so), adding a vector for counts can greatly improve the model fitting time (warmup + sampling, Supporting Information S2).

Additionally, adding counts (whether integer or non-integer) is useful for combining body size data sets that are collected with different methods (Supporting Information S2). For example, in freshwater streams, macroinvertebrates are often collected using samplers that cover 0.09 m^2. Fish in the same streams are often collected over a much larger area, such as electrofishing a 5-m wide stream for 100 m in length, yielding a sample area of $5\times100=500\,\text{m}^2$. An individual macroinvertebrate of, say, 0.01 mg and an individual fish of, say, 1000 mg would each get a count of 1 in their respective data sets. But that would not reflect their density in the food web. On a m^2 basis, the macroinvertebrate has a density of $1/0.09 = 11\,\text{m}^2$, but the fish's count (or density) should be $\frac{1}{500} = 0.002\,\text{m}^2$. The choice of units for the counts is not trivial. For example, counts per m^2 will have different confidence intervals than counts per mm^2 or per km^2 for λ (see S.1.5 in Edwards et al., 2020). We explore this in more depth in Supporting Information S2.

When body sizes are collected from the same sampling method and are not tallied or binned, all counts equal 1. If data are binned, an alternative approach, such as the `MLEbin` method, is appropriate (Edwards et al., 2020).

Converting Equation 2 into Stan allows for Bayesian estimation of λs using generalized (non)-linear mixed models. For example, an intercept-only model would look like this:

$$x_i \sim f\left(x; \lambda, x_{\min}, x_{\max}, \text{counts}_i\right) \tag{3}$$

$$\lambda = \alpha \tag{4}$$

$$\alpha \sim \text{Normal}(\mu, \sigma) \tag{5}$$

where x_i is the ith individual body size, $f\left(x; \lambda, x_{\min}, x_{\max}, \text{counts}\right)$ is the truncated Pareto distribution, λ is the size spectrum parameter (also referred to as the exponent), $x_{\min}$ and $x_{\max}$ are as described above and counts_i are the tally or density of the ith body size x in a data set. The parameter α is the intercept with a prior probability distribution. In this case, we specified a normal prior since λ is continuous and can be positive or negative.

For cases where the goal is to estimate changes in λ across space or time, the simple model above can be expanded to include predictors and/or varying intercepts and slopes:

$$x_{ij} \sim f\left(x; \lambda_j, x_{\min_j}, x_{\max_j}, \text{counts}_{ij}\right) \tag{6}$$

$$\lambda_j = \alpha + \beta z_j + \alpha_j \tag{7}$$

$$\alpha \sim \text{Normal}\left(\mu_\alpha, \sigma_\alpha\right) \tag{8}$$

$$\beta \sim \text{Normal}\left(\mu_\beta, \sigma_\beta\right) \tag{9}$$

$$\alpha_j \sim \text{Normal}\left(0, \sigma_j\right) \tag{10}$$

$$\sigma_j \sim \text{Exponential}(\phi) \tag{11}$$

where x_{ij} is the ith body size from group j. The groups might represent j sites, j experimental units or j times. The x_{ij} body sizes are distributed as a truncated Pareto with an unknown λ_j, corresponding to the size spectrum parameter for each j group, along with group specific $x_{\min_j}$ and $x_{\max_j}$. The linear model for λ_j contains an intercept α, a slope β, a continuous predictor z_j and a varying intercept for each group α_j. In this example, prior distributions are Normal for α, β and α_j. Parameters α and β require priors for their respective means μ_α and μ_β and standard deviations σ_α and σ_β. The varying intercept α_j has a mean = 0, and a standard deviation σ_j with its own Exponential hyperprior with parameter ϕ. The literature on prior choice is broad and active (Banner et al., 2020; Wesner & Pomeranz, 2021), particularly for priors on hyperparameters like σ_j (Aguilar & Bürkner, 2023; Gelman, 2006). We specify prior distributions here for clarity, but users should choose prior distributions that reflect prior knowledge. An example of checking priors with the prior predictive distribution and prior sensitivity is in the Supporting Information S4.

2.2 | Testing the models

2.2.1 | Parameter recovery from simulated data

To ensure the models could recover known parameter values, we set $j = 1, 2, \ldots, 7$ equally spaced λ values from −2.4 to −1.2. We then simulated $K = 1000$ data sets for each of the seven λs from a bounded power law using the inverse cumulative density function:

$$x_{ijk} = \begin{cases} \left(u_{ijk} x_{\max}^{(\lambda_j+1)} + \left(1-u_{ijk}\right) x_{\min}^{(\lambda_j+1)}\right)^{\frac{1}{(\lambda_j+1)}}, & \lambda \neq -1 \\ x_{\max}^{u} x_{\min}^{1-u}, & \lambda = -1 \end{cases} \tag{12}$$

where x_{ijk} is the ith individual body size from the jth value of λ_j for the kth model run (including new data simulation for each run). The variable u_{ijk} is a unique draw from a Uniform(0, 1) distribution, and all other variables are the same as defined above. We set $x_{\min} = 1$, $x_{\max} = 1000$ and simulated 300 body sizes ($i = 1,2,3, \ldots, 300$) for each j and k. To generate counts, we rounded each simulated value to the nearest 0.001 and tallied them. This generated only a small number of counts >1 (10 out of 300 body sizes). This approach was chosen to demonstrate the use of counts. For a more detailed discussion of counts, see Supporting Information S2.

Individual lambdas

After simulating the data, we estimated the λ values in three ways. First, we fit separate intercept-only models (Equation 3) to each simulated data set. This represents the common procedure of estimating λs independently before using them in later analyses (e.g. Arranz et al. (2019); Pomeranz et al. (2022)). Second, we fit a single fixed effects model of the form:

$$x_{ijk} \sim f\left(x_{jk}; \lambda_{jk}, x_{\min_{jk}}, x_{\max_{jk}}, \text{counts}_{ijk}\right) \tag{13}$$

$$\lambda_{jk} = \alpha + \beta_1 z_{1ijk} + \ \dots \ + \beta_6 z_{6ijk} \tag{14}$$

where α is the intercept representing the reference value of λ (in this case it is −2.4), β_m is the coefficient for the $m = 1, 2, \ \dots \ 6$ contrasts between the reference λ and λ_m, z_{1-6ijk} are the covariates representing the six groups containing $\lambda = \{-2.2, -2, -1.8, -1.6, -1.4, -1.2\}$. All other parameters and data are as described above. Third, we fit a varying intercepts model of the form:

$$x_{ijk} \sim f\left(x_k; \lambda_{jk}, x_{\min_{jk}}, x_{\max_{jk}}, \text{counts}_{ijk}\right) \tag{15}$$

$$\lambda_{jk} = \alpha + \alpha_{jk} \tag{16}$$

where α_{jk} are j deviations from the grand intercept α for each k simulation, with priors and hyperpriors as described in Equations 10 and 11.

The procedures above resulted in 9000 total model runs (7000 for the separate λ estimates plus 1000 each for the fixed and varying intercept models). Each model run includes newly simulated data from Equation 12. To assess how well the models captured the known λs, we estimated coverage and bias for each λ. For coverage, we generated 95% credible intervals (CrI) across each of the model runs and calculated the proportion of those CrIs that contained the known λ value. For bias, we calculated the difference between the posterior median of λ and the known λ across each model run.

Sample size and size range

We examined sensitivity to sample size (number of individual body sizes) across two λ values (−2, −1.6). For each λ, we simulated 30, 100, 300 or 1000 individuals. Each data set was fit using separate intercept-only models. We then repeated this process (data simulation and model fitting) $K = 1000$ times to estimate bias and coverage as described above.

In addition to sample size, we examined sensitivity to the size range, which can affect interpretations of λ (Sprules & Barth, 2016). To do this, we again set λ to −2 or −1.6 and then simulated $n = 300$ individuals, varying $x_{\min}$ and $x_{\max}$ so that they contained 1, 2, 3, 4 or 5 orders of magnitude in range (i.e. $x_{\min} = 1$ & $x_{\max} = 10 / 100 / 1000$, etc.). For each size range, we repeated the data simulation and model fitting 1000 times to estimate bias and coverage. We also estimated precison as the range between the lowest and highest λ estimates across each model run.

Linear variation in λ across samples

We simulated a linear regression model with a single continuous predictor z such that

$$\lambda = -1.2 - 0.05z \tag{17}$$

This contains a known intercept $\alpha = -1.2$ and a slope $\beta = -0.05$. The predictor z ranged from −1 to 1, with 10 equally spaced intervals. Values for α and β were chosen to keep λ within typical ranges of −1 to −2 across the predictors. After obtaining the 10 λ values (one for each value of z), we simulated 300 individuals from each λ using Equation 12 and setting $x_{\min} = 1$ and $x_{\max} = 1000$. The regression was fit using Equations 6–10, but without the varying intercept α_j. We repeated this procedure (data simulation and model fitting) 1000 times (see Section 2.3) and checked for parameter recovery, bias and coverage as described above.

Benefit of partial pooling and priors

Using hierarchical Bayesian models has the benefit of improving λ and regression parameter estimates with partial pooling and informative priors. These can be especially important when data from different times or places have different sample sizes. To demonstrate this, we modified the linear regression described above to include 12 values of z, one of which was an 'outlier' in which $\lambda = -1.1$ when $z = 2.5$. According to the regression equation, λ should actually equal −2.5 when $z = 2.5$. After estimating the λs, we again simulated $n = 300$ individuals from each lambda with $x_{\min} = 1$ and $x_{\max} = 1000$. However, for the outlier, we limited the number of individuals to $n = 50$. This mimics a situation in which an outlier is potentially due to a low sample size, a scenario for which partial pooling can be particularly effective (McElreath, 2020, p. 413). The purpose of this exercise is not to reflect any particular sampling scheme, but to demonstrate the importance of partial pooling and priors.

We used four techniques to estimate the relationship between z and λ. First, we used the two-step process to (1) individually estimate each lambda and (2) fit a Gaussian Bayesian linear regression between the between z and the separately estimated λs. This is akin to a no-pooling regression, in which no information about sample size or uncertainty in λ is accounted for in the λ estimates. Second, we fit the same regression but added measurement error for λ. This allows for weighting the response by the standard deviation of λs, such that the linear model has the form:

$$\lambda_j \sim \text{Normal}\left(\lambda_{\text{true},j}, \text{sd}_j\right) \tag{18}$$

$$\lambda_{\text{true},j} = \alpha + \beta z \tag{19}$$

where each λ_j has mean $\lambda_{\text{true},j}$ and standard deviation sd_j and $\lambda_{\text{true},j}$ is modelled as a linear function of z with an intercept α and slope β.

Third, we fit a linear mixed model with varying intercepts as described above, but with weak priors ($\alpha \sim N(-1.5, 1)$, $\beta \sim N(0, 0.5)$, $\sigma_j \sim \text{Exponential}(1)$). This model demonstrates partial pooling, in which the individual lambda estimates exhibit shrinkage towards the mean, particularly for the sample with 50 individuals. Additionally, the regression parameters (α, β) should be less influenced by the outlier compared to the first model. Finally, we fit a model with both varying intercepts and strong priors ($\alpha \sim N(-1.5,0.1)$, $\beta \sim N(-1,0.02)$, $\sigma_j \sim \text{Exponential}(1)$).

2.3 | Model fitting

Because the truncated Pareto pdf as described here is not available in `rstan`, we built an R package, `isdbayes` (Wesner & Pomeranz, 2023), to integrate it into `rstan` using `brms` in R (Bürkner, 2018; R Core Team, 2020). The main benefit of `brms` is that it fits Bayesian models in `rstan` using common R modelling syntax. For example, this linear regression in R, `lm(y ~ x, data=data)` becomes Bayesian using `brm(y ~ x, data=data, …)`, where `brm` will translate the model to `rstan` for MCMC sampling. The dots '…' indicate additional model specifications for the likelihood, priors, iterations, chains, etc. A short tutorial on using `isdbayes` is available at https://github.com/jswesner/isdbayes.

We specified each of the above models in `brms`, with the truncated Pareto added from the `isdbayes` package. Posteriors were explored in `rstan` (Stan Development Team, 2022) using four chains each with 2000 iterations for each model run. Two exceptions were the fixed and varying intercept models in Figure 1. For those, we specified two chains each with 2000 iterations. These values are lower than the default four chains with 2000 iterations `rstan` and `brms`, but were chosen for computational efficiency. In a separate experiment (Supporting Information S3), we re-ran a subset of those models with four chains and 2000 iterations and found no differences in the outcome. All models converged with Rhats < 1.01. Assessments of prior influence and model checking are demonstrated in Supporting Information S4. In particular, for model checking, we use simulations from the posterior predictive distributions. These simulations can check how well the model resembles the raw data. Strong deviations from raw data may indicate poor model specification or may indicate deviation from the assumption of the power law.

3 | RESULTS

3.1 | Individual lambdas

The three methods (separate models, fixed effects and varying intercepts) recaptured the true λ values (Figure 1) with no apparent evidence of bias. For example, mean bias ranged from −0.01 to 0.008, but all standard deviations included zero (Table 1). Similarly, coverage ranged from 0.93 to 0.96 with a grand mean of 0.95, indicating similarity to the nominal coverage of 0.95 (Table 1).

3.2 | Sample size and size range

Coverage ranged from 0.93 to 0.96 across sample sizes (Figure 2a), indicating good statistical coverage even at low sample sizes. However, precision increased with sample size. At $n=1000$ and $\lambda = -1.6$, the range of mean λ estimates (largest minus smallest λ) was 0.13. By comparison, it was 0.9 at $n=30$ (Figure 2a). In addition, at $n=30$, there was a slight negative bias of 0.03 units compared to the true λ (though the standard deviations all covered the true λ). This bias disappeared when $n>=100$ individuals (Figure 2).

Coverage was also consistent across size ranges, achieving nominal coverage even at size ranges of 1 order of magnitude (Figure 2b). There was also no indication of bias, with mean λ estimates ranging from −0.002 to −0.007 units away from the true λ and standard deviations including 0. However, precision was lower when body sizes ranged 1 order of magnitude (range of estimates = 0.7 and 0.6 units for $\lambda = -2$ and −1.6, respectively). Precision declined to ~0.4 and ~0.3 at body size ranges of two or more orders of magnitude and remained relatively stable (Figure 2b).

3.3 | Regression

Coverage for the intercept (α) and slope (β) parameters was 95% (Table 2, Figure 3). Bias was small for both parameters, averaging −0.0003 for α and -3e-06 for β, indicating good parameter recovery.

3.4 | Benefit of partial pooling and priors

Without partial pooling or informative priors, the two-step method was heavily influenced by the outlier, yielding a slope of −0.03 (95% CrI: −0.1 to 0.04; Figure 4a). While the credible interval contains the true slope (−0.1), there is high uncertainty in both the slope value and its sign. For example, there was only a 0.77 probability of a negative slope (and a 0.23 probability of a positive slope). Incorporating measurement error for λ made little difference, with nearly identical values of the regression parameters (Figure 4b).

Fitting the same data with a single truncated Pareto linear mixed effects model reduced the influence of the outlier, yielding a slope of −0.06 (−0.1 to 0), 50% closer to the true slope of −0.1 than the two-step model (Figure 4c). In addition, this model more reliably captured the correct sign, with a 0.96 probability of a negative slope. Adding strong priors on the slope and intercept parameters further improved the estimate (Figure 4d), with a 0.99 probability of a negative slope.

In addition to improving parameter estimates, the λ estimates themselves are improved in the partially pooled models (Figure 4c,d). For example, in the two-step method, λ in the outlier is estimated −1.1 (Figure 4a), but it is reduced to ~−1.34 with partial pooling (Figure 4c,d). Partial pooling has a minimal effect on the other λs due to their larger underlying sample size.

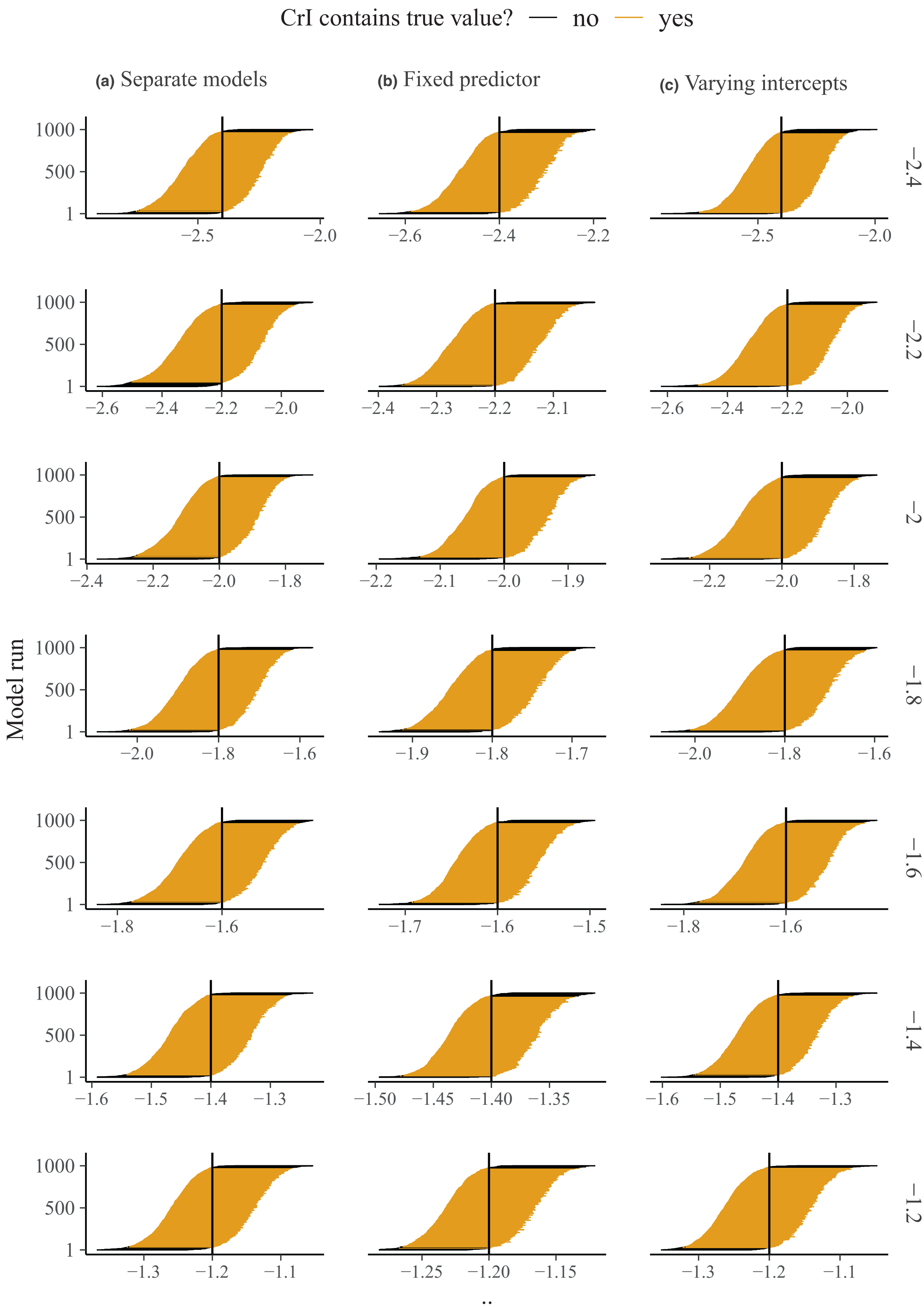
CrI contains true value? no yes
(a) Separate models
(b) Fixed predictor
(c) Varying intercepts
Model run
1000
500
1
−2.4
−2.2
−2
−1.8
−1.6
−1.4
−1.2

FIGURE 1 Modelled 95% credible intervals (CrI; $K=1000$) of seven λs using (a) separate intercept-only models for each lambda, (b) a fixed linear predictor with the λ value as a group and (c) varying intercepts. Vertical black lines show the true λ with corresponding values to the right of each row. Intervals either include the true λ (yellow) or not (black). For plotting, model runs are arranged from lowest to highest minimum value of each interval.

TABLE 1 Parameter recovery using three modelling approaches with the same data. First, separate models individually recapture known lambda values. Second, lambdas are estimated using a single fixed effects model. Third, lambdas are estimated hierarchically using a single varying intercept models. Each model and data simulation procedure is repeated 1000 times. Coverage is estimated for 95% credible intervals. Bias represents the mean and standard deviation of bias across the 1000 replicates.

True lambda	Metric	Separate models	Fixed predictor	Varying intercepts
−2.4	Coverage	0.95	0.94	0.95
−2.2	Coverage	0.93	0.96	0.96
−2.0	Coverage	0.95	0.95	0.95
−1.8	Coverage	0.95	0.94	0.95
−1.6	Coverage	0.94	0.95	0.95
−1.4	Coverage	0.95	0.95	0.94
−1.2	Coverage	0.94	0.95	0.95
−2.4	Bias	−0.006 (0.08)	0.001 (0.05)	0.01 (0.08)
−2.2	Bias	−0.011 (0.07)	−0.002 (0.04)	0.004 (0.07)
−2.0	Bias	−0.006 (0.06)	0 (0.03)	0 (0.06)
−1.8	Bias	−0.004 (0.05)	−0.001 (0.03)	−0.004 (0.05)
−1.6	Bias	−0.004 (0.04)	−0.002 (0.02)	−0.003 (0.04)
−1.4	Bias	−0.003 (0.04)	0 (0.02)	−0.005 (0.04)
−1.2	Bias	−0.001 (0.03)	0 (0.02)	−0.004 (0.03)

4 | DISCUSSION

The most important result of this work is the ability to analyse individual size distributions (ISDs) using fixed and random predictors in a hierarchical model. Our approach allows ecologists to test hypotheses about size spectra while avoiding the pitfalls of a two-step process in which λ is estimated individually for each sample and the results are then used as response variables in linear or non-linear models. The generalized mixed model with a bounded truncated Pareto merges these steps, linking the data generation process (e.g. individual body sizes) with the model predictors. This permits the use of prior probabilities and hierarchical structure on regressions of ISDs in a single analytical framework.

The ability to incorporate prior information using Bayesian updating has two practical advantages. First, adding informative prior distributions can improve model fit by limiting the MCMC sampler to reasonable sampling space. In other words, it would not be sensible to estimate the probability that λ is −1234 or − 9. Without informative priors, those values (and more extreme values) are considered equally likely and hence waste much of the algorithm's sampling effort on unlikely values (Wesner & Pomeranz, 2021).

Second, and most importantly, ecologists have much prior information on the values that λ can take. For example, global analysis of phytoplankton reveals values of −1.75, consistent with predictions based on sublinear scaling of metabolic rate with mass of −3/4 (Perkins et al., 2019). Alternatively, Sheldon's conjecture suggests that λ is −2.05 (Andersen & Beyer, 2006), a value reflecting isometric scaling of metabolic rate and mass, with support in pelagic marine food webs (Andersen & Beyer, 2006). However, benthic marine systems typically have shallower exponents (e.g. ~ −1.4; Blanchard et al. (2009)), similar to those in some freshwater stream ecosystems (~ −1.25, Pomeranz et al. (2022)). While the causes of these deviations from theoretical predictions are debated, it is clear that values of λ are restricted to a relatively narrow range between about −2.05 and −1.2. But this restriction is not known to the truncated Pareto, which has no natural lower or upper bounds on λ (White et al., 2008). As a result, a prior that places most of its probability mass on these values (e.g. $\mathrm{Normal}(-1.75, 0.2)$) seems appropriate. Such a continuous prior does not prevent findings of larger or smaller λ, but instead places properly weighted scepticism on such values.

An important assumption when setting priors is that we have a good understanding of the values that λ can reasonably take. For most of the examples here, our priors are weakly informative in the sense that they rule out clearly unreasonable values (e.g. $\lambda = -25$, etc.), but have weak effects on values within reasonable ranges (e.g.

FIGURE 2 (a) Changes in parameter estimation and coverage (numbers next to densities) as a function of sample size. Sample size is the number of individual body sizes used to estimate λ. Estimates of λ were repeated $K=1000$ times for each sample size and known λ combination. (b) The effect of size range on λ estimates. Modelled estimates ($K=1000$) of two λs using separate intercept-only models with x_{min} and x_{max} ranging one to five orders of magnitude. Horizontal black lines show the true λs (−1.6 or −2). Dots in (a and b) are the posterior median λ estimates. 95% credible intervals of those estimates (not shown for clarity) either include the true λ (yellow) or not (black).

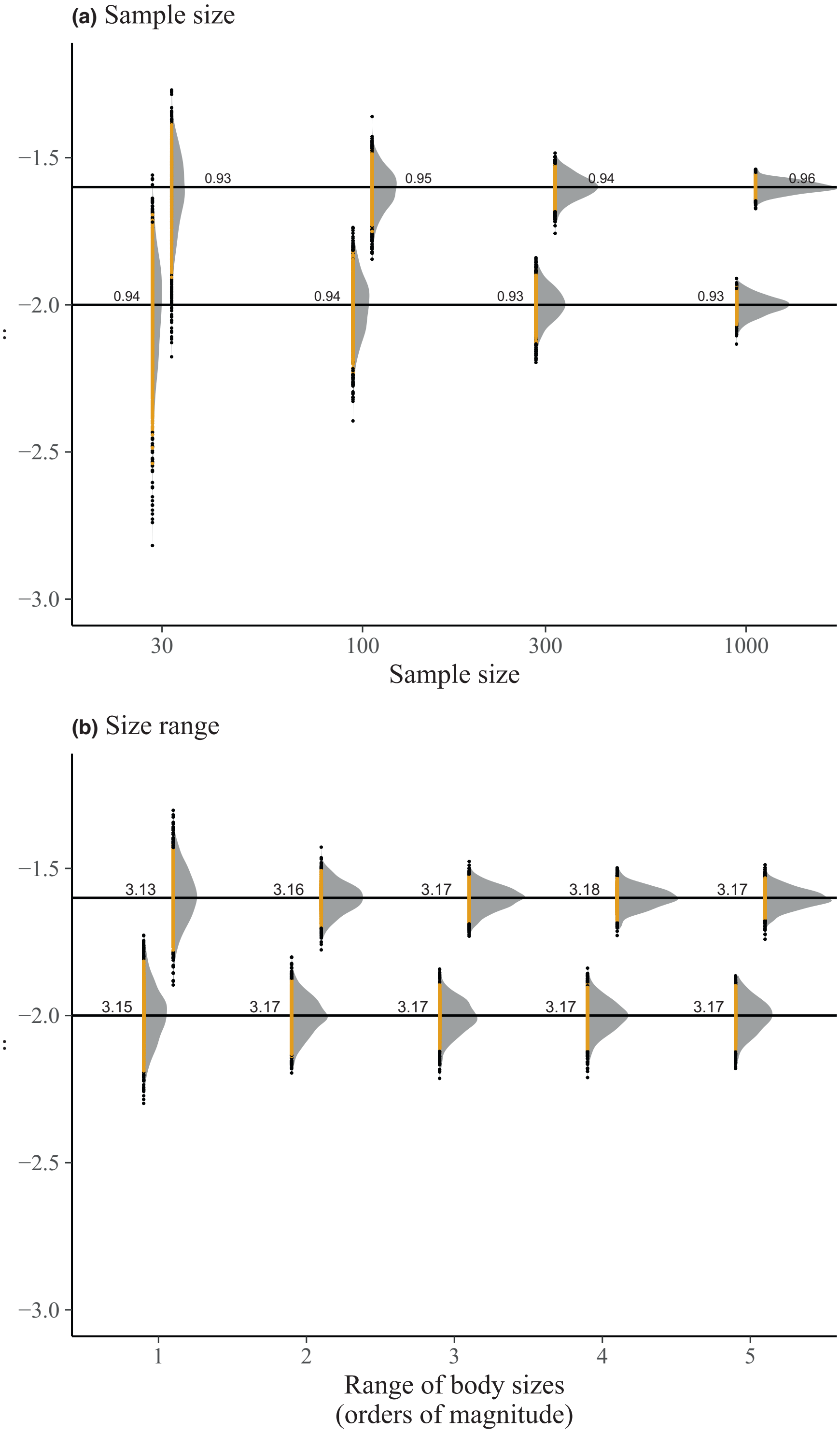

TABLE 2 Bias and 95% coverage probabilities for the intercept and slope parametres of a linear regression. Values are estimated across 1000 model runs, each of which includes simulation of body sizes and a model fit using the `isdbayes` package.

Parameter	Bias (mean, sd)	Coverage
Intercept	0 (0.01)	0.96
Slope	0 (0.01)	0.95

$\lambda \sim$-3 to 0). Most published values of λ fall into this range regardless of the method used by those studies to estimate λ (Edwards et al., 2017; White et al., 2007). However, if more informative priors are required, such as our example in Figure 4d, then caution should be used when comparing prior expectations to previously estimated λs. For example, in an analysis of marine fish trawl data, Edwards et al. (2020) found that binning methods produced λ estimates of ~−2.2 across 30 years of data. Yet reanalysis of the same data using

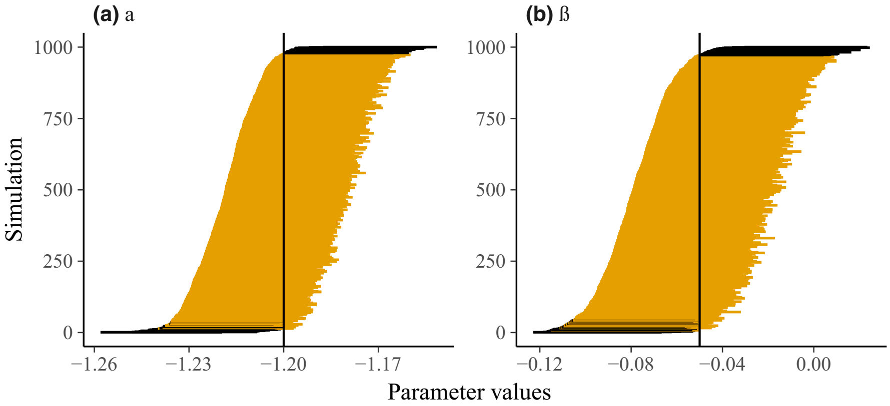

FIGURE 3 Modelled 95% credible intervals of (a) the intercept (α) and (b) the slope (β) of a generalized linear regression estimating the change in λ across a predictor. Vertical black lines show the true λ. Intervals either include the true λ (yellow) or not (black). For plotting, model runs are arranged from lowest to highest minimum value of each interval.

the truncated Pareto found λ estimates closer to −1.6. If we were to use these values to guide prior selection, then the choice of reasonable prior would clearly depend on the method used to estimate λ. The simplest approach would be to assume a fixed correction between the binning methods and the truncated Pareto when setting priors based on binning methods. Unfortunately, such a fixed correction does not appear to exist (Pomeranz et al., 2024), making it difficult or impossible to use λs from binning methods to guide informative prior selection.

Similar to priors, partial pooling from varying intercepts provides additional benefits, allowing for the incorporation of hierarchical structure and pulling λ estimates towards the global mean (Gelman, 2005; Qian et al., 2010). In the example shown here, pooling was able to downweight the influence of an outlier that had a relatively small sample size ($n = 50$ individuals compared to $n = 300$). By contrast, in the two step-method, the same outlier had a large influence on the regression outcome, because the model had no information on the number of individuals used to generate each λ. Another benefit of pooling (both from varying effects and sceptical priors) is in prediction (Gelman, 2005; Hobbs & Hooten, 2015). This becomes especially important when models are used to forecast future ecosystem conditions. Forecasts are becoming more common in ecology (Dietze et al., 2018) and are likely to be easier to test with modern long-term data sets like NEON (National Ecological Observatory Network) in which body size samples will be collected at the continental scale over at least the next 20 years (Kuhlman et al., 2016). In addition, because the effects of priors and pooling increase with smaller sample sizes, varying intercepts are likely to be particularly helpful for small samples. In other words, priors and partial pooling contain built-in scepticism of extreme values, ensuring the maxim that 'extraordinary claims require extraordinary evidence'.

One major drawback to the Bayesian modelling framework here is time. Bayesian models of even minimal complexity must be estimated with Markov Chain Monte Carlo techniques. In this study, we used the No U-Turn sampling (NUTS) algorithm (Hoffman & Gelman, 2014) via `rstan` (Stan Development Team, 2022). Stan can be substantially faster than other commonly used programs such as JAGS and WinBUGS, which rely on Gibbs sampling. For example, Stan is 10–1000 times more efficient than JAGS or WinBUGS, with the differences becoming greater as model complexity increases (Monnahan et al., 2017). In the current study, intercept-only models for individual samples with ~300 to 1500 individuals could be fit quickly (<2 s total run time (warm-up + sampling on a Lenovo T490 with 16GB RAM)). However, the hierarchical regression models took >1 h to run. These times include the fact that our models included several modifications to improve efficiency, such as weakly informative priors, standardized predictors and non-centred parameterization, each of which are known to improve convergence and reduce sampling time (McElreath, 2020). But if Bayesian inference is desired, these run times may be worth the wait. In addition, they are certain to become faster with the refinement of existing algorithms and the introduction of newer ones like Microcanonical HMC (Robnik et al., 2022).

Body size distributions in ecosystems have been studied for decades, yet comprehensive analytical approaches to testing hypotheses about them are lacking. We present a single analytical approach that takes advantage of the underlying data structures of individual body sizes (truncated Pareto distributions) while placing them in a generalized (non)-linear hierarchical modelling framework. In addition to fitting regression models, the results suggest that sample sizes >100 individuals, but optimally >1000, are sufficient to accurately estimate λ. We also found good performance at size ranges from two to five orders of magnitude, though it is important to note that this result is based on simulated data in which x_{max} is known (i.e. we 'know' it is 10, 100 or 1000 because we are using simulated data). This is more difficult in a field setting. For example, in a community with $\lambda = -1.5$, $x_{min} = 1$ g and $x_{max} = 1000$ g, there is only a 0.00016 probability of sampling an individual >999 g. In other words, if the choice of x_{max} is based only on the sample data, it is likely to underestimate the true x_{max} in the community. One approach is to set x_{max} from the largest individual caught under repeated sampling (Gjoni et al., 2024). In a field sample that ranges, say, three orders of magnitude in body size, researchers should ensure that this range reflects the likely range of true sizes in the data set. We hope that ecologists will adopt and improve on the models here to critically examine hypotheses of size spectra or other power law distributed data. Moreover, while

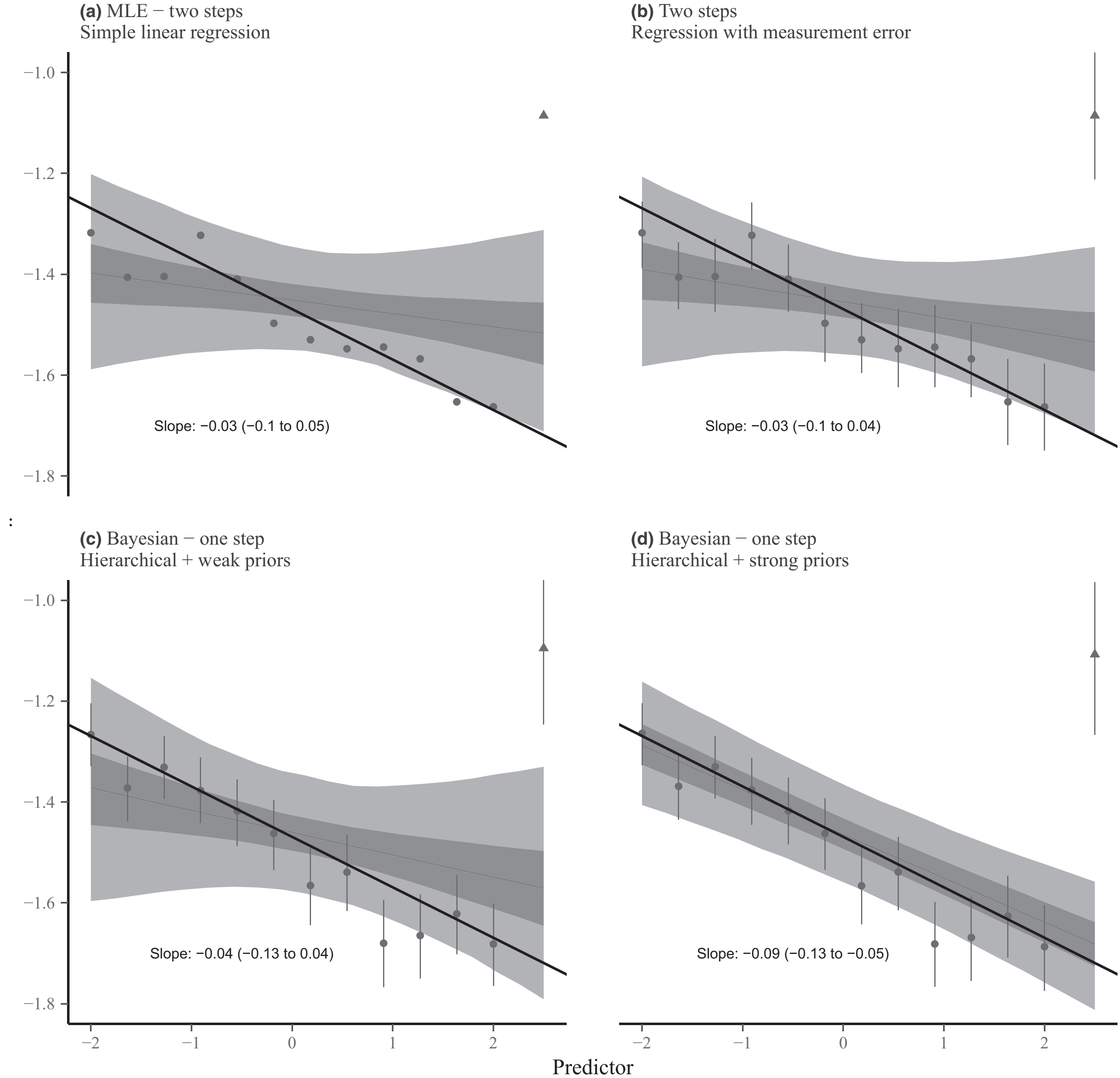

FIGURE 4 Regression results from (a) a two-step process where λs are first estimated with separate models and then used as the response variable in a Gaussian regression. (b) Same as (a), but with measurement error (posterior standard deviation of λ) included on the response variable. (c) A generalized linear mixed model with a truncated Pareto likelihood and weak priors. (d) A generalized linear mixed model with a truncated Pareto likelihood and strong priors. The solid black line shows the true regression slope. Dark shading shows the 50% CrI and light shading shows the 95% CrI. All models have the same underlying individual body size data. Points and error bars show the median and 95% CrI. For (a), only the median is shown, since the model does not include measurement error.

the examples here are for ecological size spectra, the statistical approach is not limited to ecological data, but can be applied to analysis of power law distributions that are common in a wide variety of disciplines (Aban et al., 2006; Clauset et al., 2009).

AUTHOR CONTRIBUTIONS

Jeff S. Wesner conceived the ideas and led the writing of the manuscript; Justin P. F. Pomeranz, James R. Junker and Vosjava Gjoni contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant Nos. 2106067 to JSW and 2106068 to JRJ. We thank Dr. Yuhlong Lio for statistical and mathematical advice.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14312.

DATA AVAILABILITY STATEMENT

All data and code are available https://zenodo.org/records/10699185 (Wesner, 2024).

ORCID

Jeff S. Wesner https://orcid.org/0000-0001-6058-7972
Justin P. F. Pomeranz https://orcid.org/0000-0002-3882-7666
James R. Junker https://orcid.org/0000-0001-9713-2330

REFERENCES

Aban, I. B., Meerschaert, M. M., & Panorska, A. K. (2006). Parameter estimation for the truncated pareto distribution. *Journal of the American Statistical Association*, *101*, 270–277.

Aguilar, J. E., & Bürkner, P.-C. (2023). Intuitive joint priors for bayesian linear multilevel models: The R2D2M2 prior. *Electronic Journal of Statistics*, *17*, 1711–1767.

Andersen, K. H., & Beyer, J. E. (2006). Asymptotic size determines species abundance in the marine size spectrum. *The American Naturalist*, *168*(1), 54–61. https://doi.org/10.1086/504849

Annis, J., Miller, B. J., & Palmeri, T. J. (2017). Bayesian inference with stan: A tutorial on adding custom distributions. *Behavior Research Methods*, *49*, 863–886.

Arranz, I., Hsieh, C., Mehner, T., & Brucet, S. (2019). Systematic deviations from linear size spectra of lake fish communities are correlated with predator–prey interactions and lake-use intensity. *Oikos*, *128*, 33–44.

Banner, K. M., Irvine, K. M., & Rodhouse, T. J. (2020). The use of bayesian priors in ecology: The good, the bad and the not great. *Methods in Ecology and Evolution*, *11*, 882–889.

Blanchard, J. L., Jennings, S., Law, R., Castle, M. D., McCloghrie, P., Rochet, M.-J., & Benoît, E. (2009). How does abundance scale with body size in coupled size-structured food webs? *Journal of Animal Ecology*, *78*, 270–280.

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, *85*, 1771–1789.

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*, 395–411.

Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*, 661–703.

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., Loescher, H. W., Lunch, C. K., Pijanowski, B. C., Randerson, J. T., Read, E. K., Tredennick, A. T., Vargas, R., Weathers, K. C., & White, E. P. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 1424–1432.

Edwards, A. M., Robinson, J. P. W., Blanchard, J. L., Baum, J. K., & Plank, M. J. (2020). Accounting for the bin structure of data removes bias when fitting size spectra. *Marine Ecology Progress Series*, *636*, 19–33.

Edwards, A. M., Robinson, J. P. W., Plank, M. J., Baum, J. K., & Blanchard, J. L. (2017). Testing and recommending methods for fitting size spectra to data. *Methods in Ecology and Evolution*, *8*, 57–67. https://doi.org/10.1111/2041-210X.12641

Gelman, A. (2005). Analysis of variance—Why it is more important than ever. *The Annals of Statistics*, *33*, 1–53.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian Analysis*, *1*, 515–534.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, *5*, 189–211.

Gjoni, V., Pomeranz, J. P., Junker, J. R., & Wesner, J. S. (2024). Size spectra in freshwater streams are consistent across temperature and resource supply. *bioRxiv*, 2024–01.

Goldstein, M. L., Morris, S. A., & Yen, G. G. (2004). Problems with fitting to the power-law distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, *41*, 255–258.

Hobbs, N. T., & Hooten, M. B. (2015). *Bayesian models: A statistical primer for ecologists*. Princeton University Press.

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, *15*, 1593–1623.

Jennings, S., & Blanchard, J. L. (2004). Fish abundance with no fishing: Predictions based on macroecological theory. *Journal of Animal Ecology*, *73*, 632–642.

Kerr, S. (1974). Theory of size distribution in ecological communities. *Journal of the Fisheries Board of Canada*, *31*, 1859–1862.

Kerr, S. R., & Dickie, L. M. (2001). *The biomass spectrum: A predator-prey theory of aquatic production*. Columbia University Press.

Kuhlman, M. R., Loescher, H. W., Leonard, R., Farnsworth, R., Dawson, T. E., & Kelly, E. F. (2016). A new engagement model to complete and operate the national ecological observatory network. *The Bulletin of the Ecological Society of America*, *97*, 283–287.

Martínez, A., Larrañaga, A., Miguélez, A., Yvon-Durocher, G., & Pozo, J. (2016). Land use change affects macroinvertebrate community size spectrum in streams: The case of *Pinus radiata* plantations. *Freshwater Biology*, *61*, 69–79. http://doi.wiley.com/10.1111/fwb.12680

McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Hall/CRC.

Monnahan, C. C., Thorson, J. T., & Branch, T. A. (2017). Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. *Methods in Ecology and Evolution*, *8*, 339–348.

Muller-Landau, H. C., Condit, R. S., Harms, K. E., Marks, C. O., Thomas, S. C., Bunyavejchewin, S., Chuyong, G., Co, L., Davies, S., Foster, R., Gunatilleke, S., Gunatilleke, N., Hart, T., Hubbell, S. P., Itoh, A., Kassim, A. R., Kenfack, D., LaFrankie, J. V., Lagunzad, D., … Ashton, P. (2006). Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. *Ecology Letters*, *9*, 589–602.

O'Gorman, E. J., Zhao, L., Pichler, D. E., Adams, G., Friberg, N., Rall, B. C., Seeney, A., Zhang, H., Reuman, D. C., & Woodward, G. (2017). Unexpected changes in community size structure in a natural warming experiment. *Nature Climate Change*, *7*, 659–663.

Perkins, D. M., Durance, I., Edwards, F. K., Grey, J., Hildrew, A. G., Jackson, M., Jones, J. I., Lauridsen, R. B., Layer-Dobra, K., Thompson, M. S. A., & Woodward, G. (2018). Bending the rules: Exploitation of allochthonous resources by a top-predator modifies size-abundance scaling in stream food webs. *Ecology Letters*, *21*, 1771–1780.

Perkins, D. M., Perna, A., Adrian, R., Cermeño, P., Gaedke, U., Huete-Ortega, M., White, E. P., & Yvon-Durocher, G. (2019). Energetic equivalence underpins the size structure of tree and phytoplankton communities. *Nature Communications*, *10*, 255.

Peters, R. H., & Wassenberg, K. (1983). The effect of body size on animal abundance. *Oecologia*, *60*, 89–96.
Platt, T., & Denman, K. (1977). Organisation in the pelagic ecosystem. *Helgoländer Wissenschaftliche Meeresuntersuchungen*, *30*, 575–581.
Pomeranz, J., Junker, J. R., Gjoni, V., & Wesner, J. S. (2024). Maximum likelihood outperforms binning methods for detecting differences in abundance size spectra across environmental gradients. *The Journal of Animal Ecology*, *93*, 267–280.
Pomeranz, J. P. F., Junker, J. R., & Wesner, J. S. (2022). Individual size distributions across north American streams vary with local temperature. *Global Change Biology*, *28*, 848–858.
Pomeranz, J. P. F., Warburton, H. J., & Harding, J. S. (2019). Anthropogenic mining alters macroinvertebrate size spectra in streams. *Freshwater Biology*, *64*, 81–92.
Qian, S. S., Cuffney, T. F., Alameddine, I., McMahon, G., & Reckhow, K. H. (2010). On the application of multilevel modeling in environmental and ecological studies. *Ecology*, *91*, 355–361.
R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Retrieved from https://www.R-project.org/
Reuman, D. C., Mulder, C., Raffaelli, D., & Cohen, J. E. (2008). Three allometric relations of population density to body mass: Theoretical integration and empirical tests in 149 food webs. *Ecology Letters*, *11*, 1216–1228.
Robnik, J., De Luca, G. B., Silverstein, E., & Seljak, U. (2022). Microcanonical Hamiltonian Monte Carlo. Retrieved from https://arxiv.org/abs/2212.08549
Sheldon, R., & Parsons, T. (1967). A continuous size spectrum for particulate matter in the sea. *Journal of the Fisheries Board of Canada*, *24*, 909–915.
Sprules, W. G., & Barth, L. E. (2016). Surfing the biomass size spectrum: Some remarks on history, theory, and application. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*, 477–495.
Sprules, W. G., Casselman, J. M., & Shuter, B. J. (1983). Size distribution of pelagic particles in lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, *40*, 1761–1769.
Stan Development Team. (2022). *RStan: The R interface to Stan*.
Wesner, J. (2024). *jswesner/stan_isd: Wesner et al. Bayesian hierarchical modeling of size spectra*. Retrieved from https://doi.org/10.5281/zenodo.10699185
Wesner, J., & Pomeranz, J. (2023). *Isdbayes: Bayesian hierarchical modeling of power laws using brms*.
Wesner, J. S., & Pomeranz, J. P. (2021). Choosing priors in bayesian ecological models by simulating from the prior predictive distribution. *Ecosphere*, *12*, e03739.
White, E. P., Enquist, B. J., & Green, J. L. (2008). On estimating the exponent of power-law frequency distributions. *Ecology*, *89*, 905–912.
White, E. P., Ernest, S. M., Kerkhoff, A. J., & Enquist, B. J. (2007). Relationships between body size and abundance in ecology. *Trends in Ecology & Evolution*, *22*, 323–330.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Table S1. Improvements in model run time (warmup + sampling) between long and aggregated (agg) data sets.

Table S2. First two rows of the simulated macroinvertebrate (inverts) and fish body size data.

Figure S1. One hundred posterior distributions of λ for each of four data aggregation approaches.

Figure S2. Inferences when correcting counts for sampling area are sensitive to the size of the sampling area.

Figure S3. Modeled 95% credible intervals (CrI, $K = 50$) of seven lambdas using (a) fixed effects with 2 chains, (b) fixed effects with 4 chains, (c) varying intercepts with 2 chains, or (d) varying intercepts with 4 chains.

Figure S4. Gelman-Rubin convergence diagnostics (rhats) for each parameter of $K = 50$ model runs each for (a) fixed effects with 2 or 4 chains and (b) varying intercepts with 2 or 4 chains.

Figure S5. Prior sensitivity. The data are simulated from lambda = −1.6, shown by the dotted black line.

Figure S6. One hundred simulations from (a) the prior distribution and (b) the posterior distribution after fitting the model to data.

Figure S7. Posterior predictive checks of models estimating three ISDs with true lambdas ranging from −2 to −1.3.

How to cite this article: Wesner, J. S., Pomeranz, J. P. F., Junker, J. R., & Gjoni, V. (2024). Bayesian hierarchical modelling of size spectra. *Methods in Ecology and Evolution*, *15*, 856–867. https://doi.org/10.1111/2041-210X.14312