**Supplementary Information File for**

# *Size spectra in freshwater streams are consistent across temperature and resource supply*

Vojsava Gjoni[1]

Justin P.F. Pomeranz[2]

James R. Junker[3]

Jeff Wesner[1]

[1]University of South Dakota, Department of Biology, Vermillion, SD 57069

[2]Colorado Mesa University, Department of Physical and Environmental Sciences, Grand Junction, CO 81501

[3]University of North Texas, Department of Biological Sciences, Denton, TX 76203

*corresponding author: vojsava.gjoni@usd.edu

***Model Checking***

We used examined model fit using Bayesian ($R^2$)[1], posterior predictive checks [2], and Bayesian P-values[3]. Each of these measures summarize comparisons between the raw data and data predicted from the posterior distribution of the fitted model. To do this, we first needed to remove the $counts$ variable from the raw data. It contains the density of each individual body size, which allows us to combine the fish and macroinvertebrate data sets while accounting for the different collection areas and relative abundances of each taxa[4]. However, while the densities are included in the likelihood when fitting the model, they are not included in the random number generator for simulating data. Therefore, to remove them, we re-sampled 5,000 individual body sizes with replacement from each of the 133 samples, weighting each sample by its density in no/$m^2$. This generated a vector of individual body sizes, each with an implied density of 1.

To simulate new data from the posterior, we first extracted the posterior distribution of $\lambda$ for each of the $j = 133$ samples using the add_epred_draws() function from the *tidyabayes* package (Kay 2023). This function applies the following:

$$\lambda_j^k = g\left(\boldsymbol{\theta}_j^k, \mathrm{X}_j\right)$$

where $\lambda_j^k$ is the $k^{\text{th}}$ posterior draw from sample $j$, derived from the linear equation containing the $k^{\text{th}}$ parameter values $\boldsymbol{\theta}$ and data X associated with sample $j$. From the first 100 $k$ draws of each $\lambda_j$, we simulated 5,000 individual body sizes using the inverse cumulative density function[5] via the *rparetocounts* function from the R package *isdbayes*[4].

The end result is 5000 simulated individual body sizes from each of the 133 NEON samples. This allowed us to compare model fit at the sample level. We also compared fits from the full model using the posterior mean estimate of $\lambda$. In other words, we simulated the full data set rather than data sets for each sample $j$.

To determine how well the model recaptures the raw data, we visually compared the simulated data, $y_{pred,j}$, and the raw re-sampled data, $y_j$ of each $j$ sample. We also calculated the geometric mean for each $j$ prediction and raw data. We then calculated a Bayesian P-value as the proportion of posterior

draws for which the geometric mean was greater than the raw value. Proportions >0.1 or <0.9 are generally indicative of poor model fit, indicating a mismatch in the $ypred$ and $y$[3].

We calculated a Bayesian $R^2$ using the following formula[1]:

$$R^2 = \frac{\mathrm{V}_{ypred}}{\mathrm{V}_{ypred} + \mathrm{V}_{res}},$$

where $\mathrm{V}_{ypred}$ is the variance of $ypred$ and $\mathrm{V}_{res}$ is the variance of the residuals $ypred - y$. We repeated this equation for each of 1,000 $k$ draws from the posterior, generating a distribution of $R^2$.

Finally, to visualize the size spectrum, we plotted the cumulative distribution function from the fitted model against the raw data[6]. In particular, we obtained the posterior median and 95% CrI of $\lambda$ for each of the 133 body size samples. From those, we calculated the following cumulative distribution function:

$$\mathrm{P}(X \geq x) = \frac{1 - \left(x^{\lambda+1} - \left(\mathrm{xmin}^{\lambda+1}\right)\right)}{\left(\mathrm{xmax}^{\lambda+1} - \mathrm{xmin}^{\lambda+1}\right)}$$

where $\mathrm{P}(X \geq x)$ is the probability of obtaining a body size $X$ that is greater than or equal to a give size $x$ in the data set. The largest individual in the data set has $\mathrm{P}(X \geq x) = 1$. The smallest individual has $\mathrm{P}(X \geq x) = 0$, and all other individuals are in between. To plot the raw data on top of this function, we ranked body sizes within samples from 1 to 5000 (largest to smallest). We then multiplied $\mathrm{P}(X \geq x)$ by 5000 so that the y-axis contains the number of individuals $\geq x$, rather that a probability *per se*. Because the raw re-sampled data are unlikely to contain the true $x_{max}$ or $x_{min}$ we used the $x_{max}$ and $x_{min}$ of the re-sampled data in the calculation for the cumulative distribution function.

*Results*

The model had a Bayesian $R^2$ of $0.47 \pm 0.02$ (mean $\pm$ sd), indicating good fit explaining ~47% of the variance of new data. Posterior predictive checks revealed generally good fit (Figure ED4). First, the predictive distributions in Figure ED4 generally resemble the raw re-sampled distributions. This suggests that the truncated Pareto is a reasonable likelihood for these data. However, there are several

clear discrepancies. For example, there is variation in fit among sites, with MCDI and KING sites appearing strong, while sites like GUIL and BLUE have more of their raw distributions in larger body size ranges compared to the posterior predictive. These may indicate either deviation from a power law at these sites due to underlying mechanisms in the food web. Alternatively, they can indicate an under sampling of either large or small individuals during field collections, or a combination of the two. Our perspective is that the truncated Pareto provides a reasonably good fit to the data overall.

Despite some variation among sites between $y_{pred}$ and $y$, there is strong agreement in the geometric means (GM) across samples. Figure ED5 shows the GM($y_{pred}$) compared to the GM($y$) across all 133 samples. The GM($y$) is consistently within the 95% credible intervals of GM($y_{pred}$) at each sample (Figure ED5). Moreover, Bayesian P-values across all samples ranged from 0.13 to 0.84 with a mean of 0.5 and sd of 0.14, again indicating good overall model fit.

**Data Sources**

We used six data sources collected by the National Ecological Data Observatory (Extended Data Table 2): Macroinvertebrates[7], Fish[8], Temperature[9], Stream Discharge[10], Oxygen[11], and Organic Matter, which was measure directly from samples maintained at the NEON Biorepository[12].

**References**


1. Gelman, A., Goodrich, B., Gabry, J. & Vehtari, A. R-squared for Bayesian Regression Models. *The American Statistician* **73**, 307–309 (2019).
2. Gabry, J., Simpson, D., Vehtari, A., Betancourt, M. & Gelman, A. Visualization in Bayesian Workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society* **182**, 389–402 (2019).
3. Hooten, M. B. & Hobbs, N. T. A guide to Bayesian model selection for ecologists. *Ecological Monographs* **85**, 3–28 (2015).
4. Wesner, J. S. & Pomeranz, J. Isdbayes: Bayesian Hierarchical Modeling of Power Laws Using Brms. (2023).
5. Wesner, J. S., Pomeranz, J. P. F., Junker, J. R. & Gjoni, V. Bayesian hierarchical modelling of size spectra. *Methods in Ecology and Evolution* **n/a**,.
6. Edwards, Andrew. sizeSpectra: Fitting Size Spectra to Ecological Data Using Maximum Likelihood. (2023).
7. NEON. NEON 2023 (National Ecological Observatory Network). Macroinvertebrate collection (DP1.20120.001), RELEASE-2024. (2023).
8. NEON. NEON 2022 (National Ecological Observatory Network). Fish electrofishing, gill netting, and fyke netting counts (DP1.20107.001). (2022).
9. NEON. NEON 2023 (National Ecological Observatory Network). Temperature (PRT) in

surface water (DP1.20053.001). (2023).
10. NEON. NEON. 2023. Continuous discharge (DP4.00130.001). (2023).
11. NEON. NEON. 2023. Water quality (DP1.20288.001). (2023).
12. Yule, Gilbert & Franz. Designing Biorepositories to Monitor Ecological and Evolutionary Responses to Change. (2020) doi:10.5281/zenodo.3880411.