Extended Data: 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

Model Checking

We examined model fit using Bayesian R^2 (Gelman et al. 2019), posterior predictive checks (Gabry et al. 2019), and Bayesian P-values (Hobbs and Hooten 2015). 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 (J. Wesner and Pomeranz 2023). 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 λ for each of the $j = 133$ samples using the `add_epred_draws()` function from the `tidyabayes` package (Kay, n.d.). This function applies the following:

$$\lambda_j^k = g(\theta_j^k, X_j)$$

where λ_j^k is the kth posterior draw from sample j, derived from the linear equation containing the kth parameter values θ and data X associated with sample j. From the first 100 k draws of each λ_j, we simulated 5,000 individual body sizes using the inverse cumulative density function J. S. Wesner et al. (2023) via the `rparetocounts` function from the R package `isdbayes` J. Wesner and Pomeranz (2023).

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 λ. 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 (Hobbs and Hooten 2015).

Finally, we calculated a Bayesian R^2 using the formula given in Gelman et al. (2019):

$$R^2 = \frac{V_{ypred}}{V_{ypred} + V_{res}},$$

where V_{ypred} is the variance of $ypred$ and 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.

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 ED1). First, the predictive distributions in Figure ED1 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 deviations 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 ED2 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 ED2). 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.

References

Gabry, J., D. Simpson, A. Vehtari, M. Betancourt, and A. Gelman. 2019. "Visualization in Bayesian Workflow." *J. R. Stat. Soc. A. Stat* 182: 389402.

Gelman, Andrew, Ben Goodrich, Jonah Gabry, and Aki Vehtari. 2019. "R-Squared for Bayesian Regression Models." *The American Statistician*.

Hobbs, N. Thompson, and Mevin B. Hooten. 2015. *Bayesian Models: A Statistical Primer for Ecologists*. Princeton University Press. https://doi.org/10.23943/princeton/9780691159287.001.0001.

Kay, Matthew. n.d. *tidybayes: Tidy Data and Geoms for Bayesian Models*. https://doi.org/10. 5281/zenodo.1308151.

Wesner, Jeff S, Justin PF Pomeranz, James R Junker, and Vojsava Gjoni. 2023. "Bayesian Hierarchical Modeling of Size Spectra." *bioRxiv*, 2023–02.

Wesner, Jeff, and Justin Pomeranz. 2023. *Isdbayes: Bayesian Hierarchical Modeling of Power Laws Using Brms*.
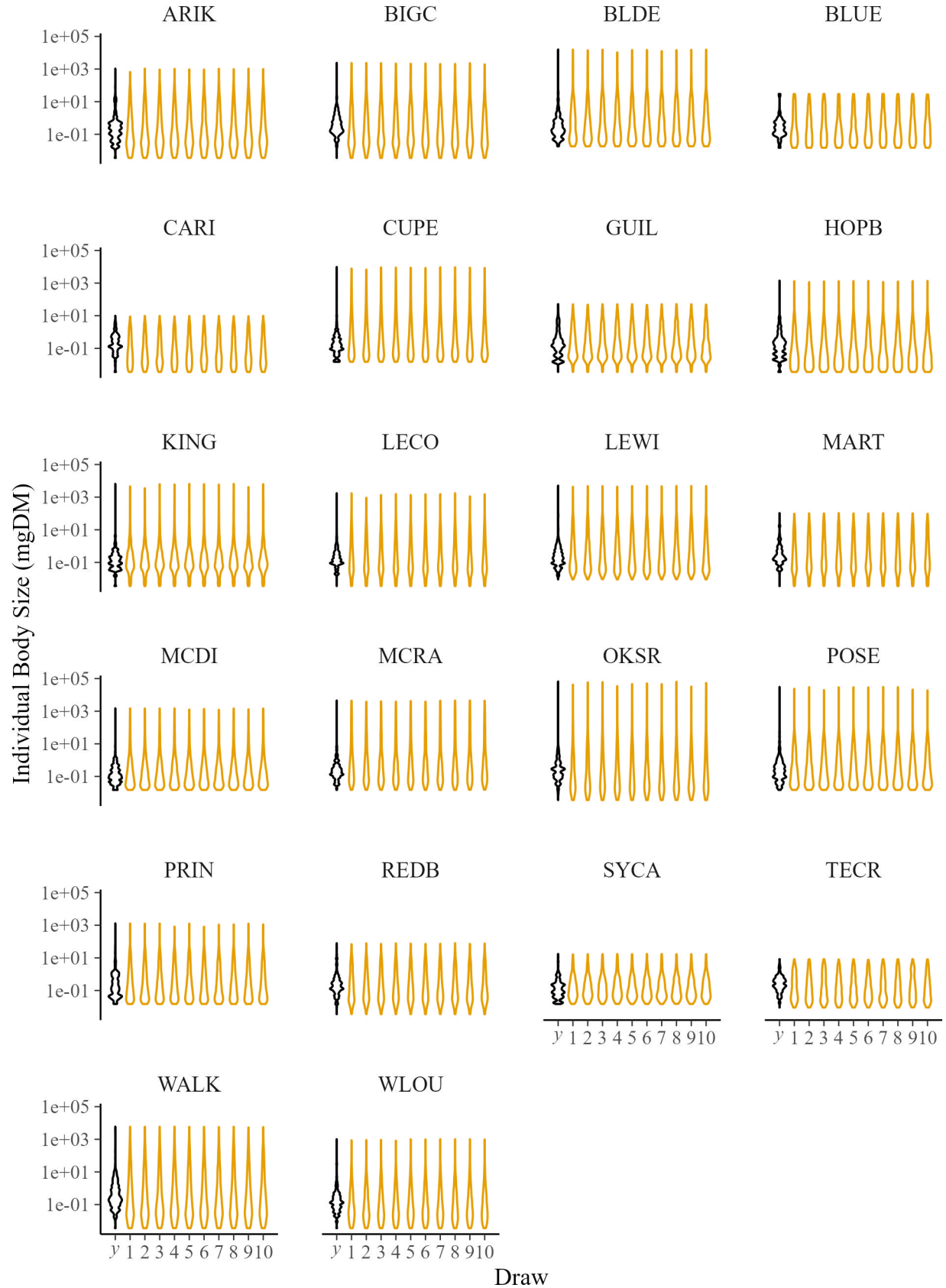
Figure S 1: Posterior predictive check showing the distribution of the raw re-sampled (y) data compared to 10 draws from teh posterior predictive distribution. The results are visualized at the site level here, but the distributions reflect simulations from each individual sample within a site.
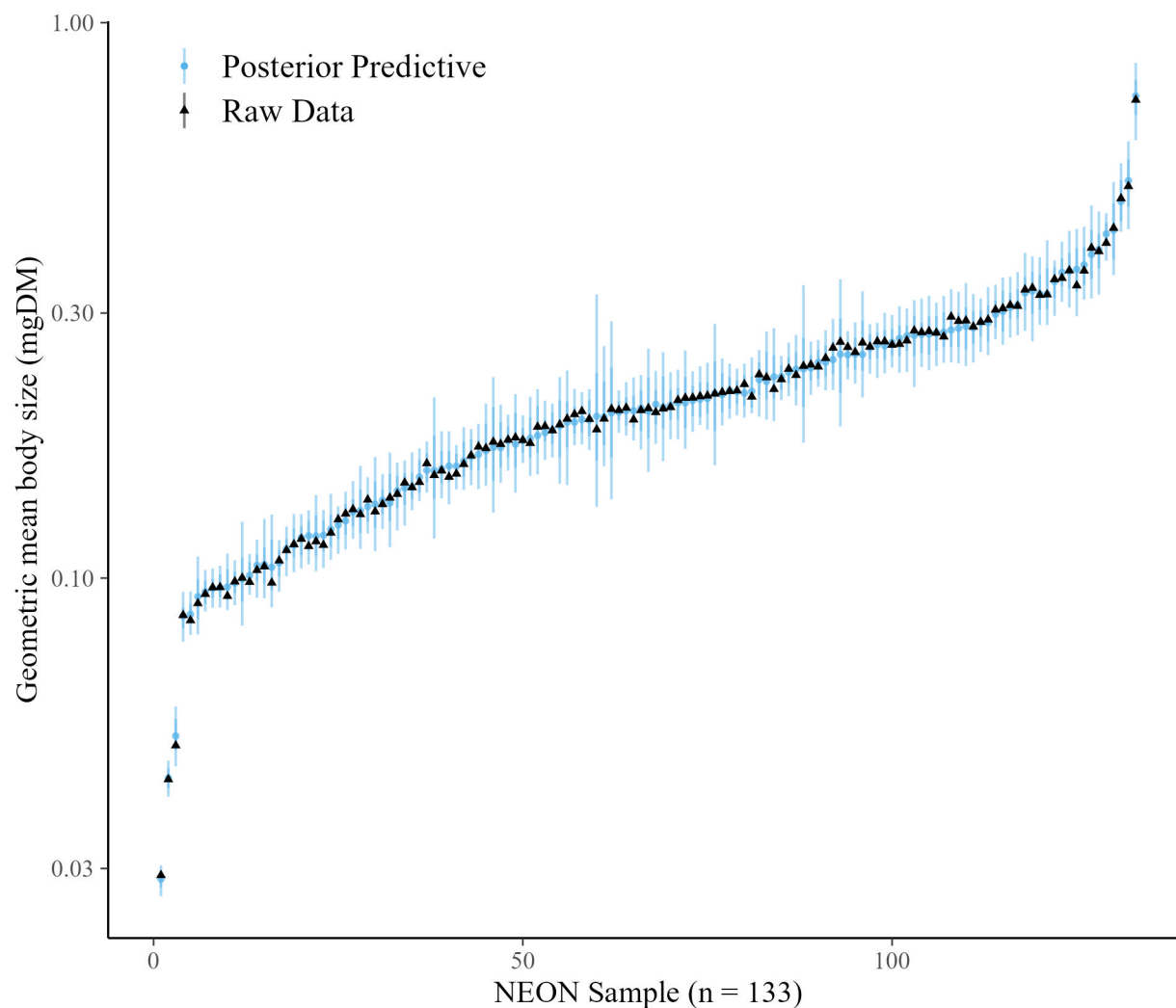
Figure S 2: Posterior predictive check showing the geometric mean body size calculated from the raw re-sampled data (black triangle) compared to the geometric mean calculatd from the posterior predictive distribution (blue dot and line). The blue dot and line represent the median and 95% quantiles from 100 draws of the posterior distribution