

## **Mid-Project Update**

Kasturi Bhatt (1009090877) and Anna Whitehouse (1009011837)  
Department of Ecology and Evolutionary Biology, University of Toronto  
EEB313: Quantitative Methods in R for Biology  
Metu Yuksel and Zoë Humphries  
November 21, 2024

## Mid-Project Update

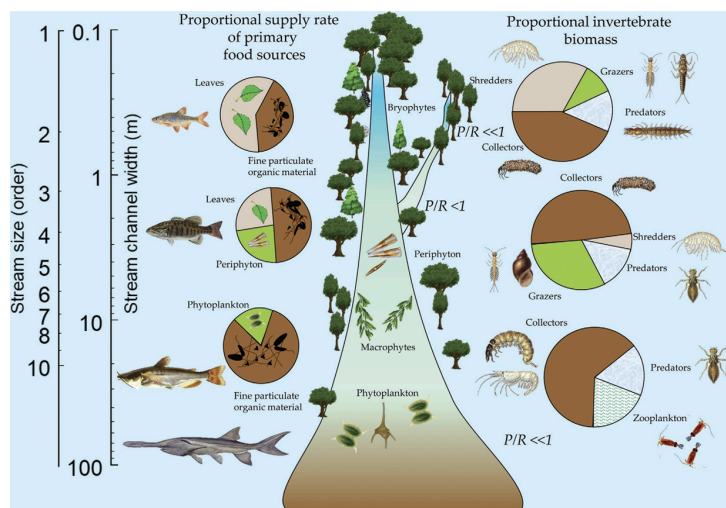
### Hypotheses and Predictions

#### Hypotheses:

- $H_0$ : There is no change in macroinvertebrate group abundance along river gradients.
- $H_1$ : There is a statistically significant change in macroinvertebrate group abundance along river gradients.

#### Predictions:

Abundances of functional feeding groups will change from headwater downstream, in relation to the expected abundances in the River Continuum Concept (see Figure 1).

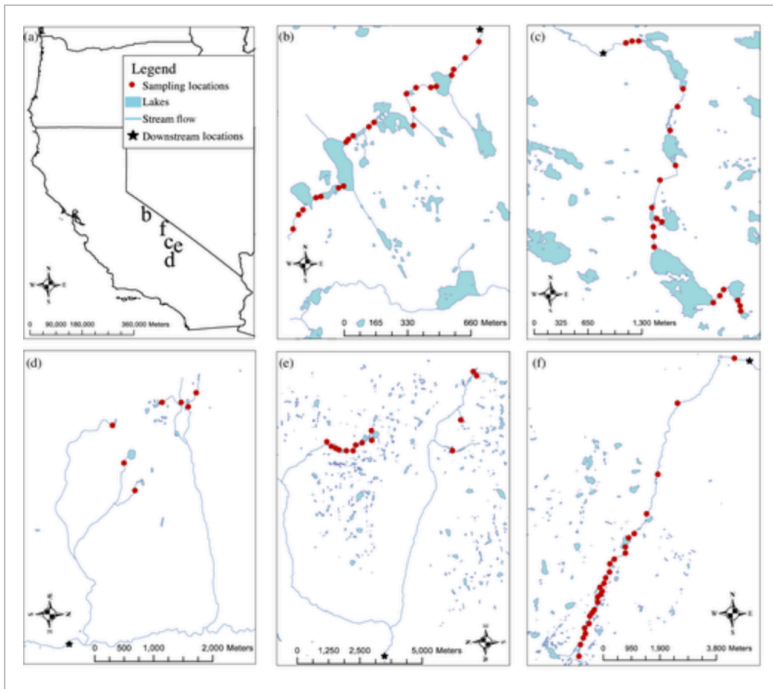


**Fig.1 - River Continuum Concept.** Source: Dodds and Maasri, 2022

### About our Data

#### Collection Methods:

The researchers sampled five distinct lake-stream networks in Sierra-Nevada, California. They sampled along spatial gradients within each network. They took multiple trials of macroinvertebrate abundance counts at each sampling location. This included identifying species, counting species abundances, and mapping the sampling sites along the networks (see Figure 2).



**Fig.2 - Map of the 5 stream-river networks and sampling sites.** Source: Green et.al, 2022a.

#### Data Wrangling:

First, we took the dataset from Green et al. (2022b), and transposed it in Excel to pivot from wide to long data, so that species and genus names were in a column rather than a row. Using family information from macroinvertebrates.org (n.d.), and the Benthic Macroinvertebrate Master Taxa List (“Benthic macroinvertebrate master taxa list”, n.d.) we added a family to every species or genus that was in the original count data from Green et al. (2022b). We then used the Benthic Macroinvertebrate Master Taxa List as a reference table to add in functional feeding group information based on family. We then removed species, genus and family names, and kept only counts by functional feeding group and site.

To simplify the data, we took out observations from river systems which had branched patterns, as this would not constitute a river continuum. We kept the Cascade Lake Network, the Evolution Lake Network and the Rock Creek Lake Network, as they were all unbranched rivers. We also filtered out UN (“unidentified”) and PI (“piercers”) functional feeding groups, because they did not fit with the functional feeding groups described in the River Continuum Concept. We did this confidently, by looking at the relative proportions of feeding groups, and by finding that piercers do not make up much of the sample. We felt that it was acceptable to exclude them.

In order to see clearer patterns and to better link our data to the River Continuum Concept, we simplified feeding groups into five groups (see Table 1), so that we could re-group the feeding groups which act in multiple categories. Note that some original feeding groups fit into multiple broad feeding groups, and so they were counted in both categories.

<u>broad feeding group</u>	<u>original feeding group</u>
CG: Collector-grazers	<ul style="list-style-type: none"> <li>- CG: Collector-grazer</li> <li>- CG/PR: Collector-grazer and predator</li> <li>- CG/SC: Collector-grazer and scraper</li> </ul>
CF: Collector-filterers	<ul style="list-style-type: none"> <li>- CF: Collector-filterer</li> </ul>
SC: Scrapers	<ul style="list-style-type: none"> <li>- SC: Scraper</li> <li>- CG/SC: Collector-grazer and scraper</li> </ul>
SH: Shredders	<ul style="list-style-type: none"> <li>- SH: Shredder</li> </ul>
PR: Predators	<ul style="list-style-type: none"> <li>- PR: Predator</li> <li>- CG/PR: Collector-grazer and predator</li> </ul>

**Table 1 - New functional feeding group categorizations.**

We also calculated the mean count per functional group, by taking the mean of trials at each site. Finally, we calculated proportions of each functional feeding group at each site, in relation to its total count in the whole river system.

### **Exploratory Data Analysis**

We plotted the site against mean proportion, grouped by functional feeding group, and faceted by the river system. This allowed us to see how the proportions of different functional feeding groups changed along the river gradient. Where necessary, we plotted twice, where we removed groups with high abundances in the second plot, in order to see the patterns in the groups with lower abundances.

### **Analysis Plan**

A multinomial distribution is used to determine the likelihood of a set of outcomes. It allows us to see sites as sides on a dice, where a “success” is a singular count of a functional feeding group in a site. We plan to use the multinomial distribution (specifically, the function `dmultinom()`) to predict the probability that we observe the given abundance of a functional group at a given site. We will do this for each river and each functional group independently. This will allow us to compare our data to see if the change in probabilities follows the general pattern that we expect to see given the River Continuum Concept. We will do this by looking at the variance in the distribution at each site, and assessing whether the values are significantly different between sites.

The assumptions of a multinomial distribution include: (i) fixed number of trials; (ii) independent trials; (iii) mutually exclusive outcomes in each trial, and; (iv) that each outcome

occurs with a given probability (“2.3 - The Multinomial Distribution”, n.d.). From the data, we know that there are a fixed number of trials (being the total count of a functional feeding group along a river) and that the trials are independent, as each organism is sampled independently. The outcomes are mutually exclusive, given the fact that the presence of a functional group at one site does not preclude it from being at another. Finally, each outcome occurs with a given probability, which is equal to the proportion that an abundance at one site makes up in relation to the total count of that functional feeding group in the whole river system.

## References

2.3—*The Multinomial Distribution*. (n.d.). STAT504 | Analysis of Discrete Data; Penn State: Emberly College of Science. <https://online.stat.psu.edu/stat504/lesson/2/2.3>

Dodds W, Maasri A. 2022. The River Continuum Concept. In: Encyclopedia of Inland Waters. Vol. 2. 2nd ed. Elsevier. p. 237–243.

*Benthic macroinvertebrate master taxa list | bionet*. (n.d.). Retrieved November 19, 2024, from <https://programs.iowadnr.gov/bionet/Inverts/Taxa>

Green, MD, Anderson KE, Herbst DB, Spasojevic MJ. 2022a. Rethinking biodiversity patterns and processes in stream ecosystems. *Ecological Monographs*. 92(3). doi:10.1002/ecm.1520

Green MD, Anderson KE, Herbst DB, Spasojevic, MJ. 2022b. Species and environmental datasets from Sierra Nevada, CA (USA) streams in lake-stream networks [Dataset]. Dryad. <https://doi.org/10.5061/dryad.2fqz612qw>

*Macroinvertebrates.org*. (n.d.). Retrieved November 19, 2024, from <https://www.macroinvertebrates.org/>