Weighted Averaging

N-STEPS Objectives

Provide regions, states, and tribes with support related to nutrient criteria development

Provide access to expert assistance with issues related to nutrient criteria development and implementation

Improve communication nationwide.

What Is It?

Weighted averaging is a method for computing an average where some values receive greater weight than others. This technique is used to identify chemical (including nutrient) optima for algal taxa and was developed for use in paleolimnology to reconstruct historic environmental conditions in lakes based on which algal taxa are present in lake sediments. Because organisms vary in their competitive abilities under different conditions, their preferences or environmental optima vary. But organisms also integrate the average environmental conditions over long periods of time, during which water chemistry may be highly variable. Use of algal composition, combined with information about algal optima is, therefore, a strong approach to identifying the true conditions of a waterbody when those conditions are difficult to characterize. These optima also reflect which algal taxa prefer high or low nutrient concentration and, therefore, which are good indicators of excess nutrients.

Example Question: Which algal taxa prefer high nutrient concentrations?

How is it Applied to Nutrient Criteria Development?

Nutrient criteria development involves three main processes: identifying relationships between biological responses and nutrient stressors, examining these relationships, and establishing nutrient and/or biological thresholds or criteria.

Weighted averaging is being used to identify nutrient optima for different algal taxa. Nutrient optima are part of the autoecological information about algal taxa that is important for applying them to nutrient criteria development. Because taxa differ in their optima, the algal composition of a site at anytime provides a fingerprint of the environmental conditions. This fingerprint can be used to infer the true nutrient conditions. Algal optima also allow one to identify which taxa indicate high nutrient or low nutrient conditions.

How Does It Work?

Weighted averaging simply calculates the average nutrient concentration for sites where a taxon is found weighted by the relative abundance of that taxon at each of those sites. The formula is:

$$Weighted\ Average = \frac{\sum_{i=1}^{N} x_i w_i}{\sum_{i=1}^{N} w_i}$$

Where x_i is the nutrient concentration at site i, and w_i is the relative abundance of a taxon at site i. This value is summed across all sites where a taxon is found and divided by the total sum of relativel abundances. The result is the average total phosphorus concentration across sites where the taxon is found weighted by relative abundance and is an estimate of the nutrient optimum for that taxon. Note that this is optima also allows one to identify which taxa prefer relatively high vs low nutrient concentrations and, therefore, indicator taxa for enrichment.

Nutrient concentrations can then be estimate for any sites based on its taxonomic composition. Algal taxonomic abundance data for the site is combined with the optima values for all the taxa. The inference models often have to be corrected and the most common approach is to use weighted averaging partial least squares (WAPLS) regression. Models can be validated by comparing estimates across many sites with actual nutrient concentrations measured for those sites.

Data Requirements

Weighted averaging requires algal relative abundance data and nutrient concentration data across a large gradient in nutrient concentrations. It is best to have as large a gradient as possible and the most data as possible.

What Should You Look For & Report?

Essentially, weighted averaging generates a nutrient optimum for each taxon. These should be reported, ideally with the sample size. Plots of relative abundance against nutrient concentrations can be reported as well, so users can evaluate the variability in the data.

Inference modeling involves an entire additional set of analyses and output and may be the focus of a future fact sheet. Users are encouraged to see the literature below for more information on inference modeling.

Pros

- Effective technique for identifying taxa optima
- Can be used to infer true average nutrient concentrations at a site
- Can be used to identify taxa that indicate enrichment

Cons

- Require substantial data
- Can be biased, depending on sample size

Alternatives

Maximum likelihood estimation

Citations

Potapova, M.; Charles, D. F.; Ponader K. C.; Winter D. M. 2004. Quantifying Species Indicator Values for Trophic Diatom Indices: A Comparison of Approaches. Hydrobiologia, 517: 25-41.

Ponader, K.; Charles, D. F. and T.J. Belton. 2005. Diatom-based TP and TN inference models and indices for monitoring nutrient enrichment of New Jersey streams. Ecological Indicators 7: 79-93

Sokal, R.R. and F.J. Rohlf. 1995. Biometry 3rd Edition. Freeman, New York.

ter Braak, C. J. F.; Juggins, S. 1993. Weighted Averaging Partial Least Squares Regression (WA-PLS): An Improved Method For Reconstructing Environmental Variables from Species Assemblages. Hydrobiologia, 269/270:485-502.

Winter, J.G., and Duthie, H.C. 2000. Epilithic diatoms as indicators of stream total N and total P concentrations: Journal of the North American Benthological Society, 19: 32-49.