

Analyzing and Interpreting Aquatic Biological Data (Newsletter)*

January 10, 2014

Regulatory and resource agencies sometimes require collection of benthic macroinvertebrates and/or fish for baseline data or permit compliance. It is too common for them to not know what they will do with the data. Correctly analyzing and interpreting these data yields valuable information that operators and regulators can use to make well-informed decisions regarding Clean Water Act compliance.

Aquatic biotic communities reflect ambient water quality conditions much better than do chemical concentrations. This is particularly true for benthic macroinvertebrates as they are less mobile than fish. Macroinvertebrates also quickly recolonize areas from which they were extirpated, frequently within a month.

Selecting appropriate statistical models requires understanding the differences between biological and physico-chemical data. For example, the latter are continuous (infinite values between any two end points such as a degree of temperature or mg/L of concentration) while the former are counts. Count data are not represented by continuous frequency distributions such as the normal (bell) curve, but the logarithms of counts might be normally distributed. The absence of an organism at one or more locations might be statistically significant; then again, it might not be: the organism might be present but not collected. In this situation the Bayesian approach to statistical models is more appropriate than is either the frequentist or maximum likelihood approach. The interpretation of model results needs to be carefully considered for the specific question that needs to be answered.

Traditional characterization of these biota used indices based on taxonomic identification to species. Even with mixed taxonomic levels the indices are applied and arbitrary assumptions of thresholds between "good" and "not good" used to interpret the results. There are much more robust analyses available, particularly those based on categories of feeding behaviors (shredding, filtering, gathering, grazing, predating) rather than taxonomic identification below the family level. Selection of statistical models are based on specific questions to be answered.

*Copyright ©2014 Applied Ecosystem Services, Inc.

Characterize each data set by drawing boxplots of the relative proportions of each functional feeding group (FFG) and by calculating the Bray-Curtis dissimilarity distances among collections by location or time. The distribution of FFG proportions reflect natural variability and the Bray-Curtis distances indicate when patterns change. The results of this characterization can be applied with results of other statistical models to establish water quality standards specific to each stream network or specific reaches within a larger river system.

Federal and state regulators consider 'aquatic life' to be the highest and best designated beneficial use; if that use is attained so are other beneficiaries such as wildlife, cattle, and irrigation. Aquatic biota, by definition, are aquatic life. Therefore, demonstrating the range of inherent natural variability of biotic community structure. Applying appropriate regression, time series, and classification models to these data reveals those spatial, temporal, physical, and chemical factors that result in the observed biotic community structure. If industrial activities do not change these patterns or effects of explanatory variables in a statistically significant way, then regulators and the public can be confident that the waters have attained their highest and best beneficial use.

—
All newsletters, white papers, and other technical resources can be freely downloaded from <http://www.appl-ecosys.com/publications/>.