Making Sense of Complex Environmental Biological Data (Newsletter)*

July 19, 2013

From baseline conditions for environmental impact assessments to compliance with regulatory permit conditions regulated companies collect biological data and report analytical results to regulators and other interested parties.

Historically, analyses used biotic diversity and integrity indices. These attempt to summarize highly complex natural ecosystems in a single number believed to make comparisons and decisions easier. While these indices are based on ecological theory they are very difficult, even impossible, to measure and quantitatively compare. Diversity indices commonly contain different taxonomic levels (species, genus, family, order) and assign arbitrary thresholds between "good" and "not-good." The difference between 1 and 2 may not be the same as the difference between 2 and 3.

Biotic integrity indices are constructed specifically for each state, major river basin, or local area, and "integrity" (often expressed as "poor," "fair," "good," or "excellent") is subjective and cannot be measured like distance or time. With these indices there is no single, consistent process producing objective and meaningful site-specific results that can be used everywhere. Also, the question each index is intended to answer is not always clearly defined and it is common for there to be uncertainty how to use these numbers in making regulatory decisions.

The good news is that alternative, highly robust analytical approaches are available and offer many advantages to operators and regulators.

Many robust statistical models will answer questions asked of biological data. Are there changes over time at the same location and are there differences among locations? Can observed changes be attributed to operations and not inherent natural variability? The specific concerns determine whether frequentist, maximum likelihood, or Bayesian statistical models are used. All models are based on solid mathematical foundations and produce results that are technically sound and legally defensible.

These statistical models can be used with count data (e.g., number of fish in a stream, density of plant species in an area), summary data (e.g., species richness), and functional data (e.g., benthic macroinvertebrate feeding strategies; breeding make-to-female ratios in wildlife populations). These models can also identify factors contributing to the observed results; for example, associating benthic macroinvertebrate taxa with water chemistry and predicting environmental conditions based on the observed taxa. These models can also answer project- and site-specific questions of very high complexity incorporating economic, social, and political factors as well as biotic data.

Equally important, statistical model results can be presented as plots or graphs which make relationships and results very easy to communicate to non-technical decision-makers, stake holders, and the public. Effective communication of the meaning and significance of biotic data helps avoid paralysis by analysis and facilitates understanding of complex systems by everyone involved in the process.

Statistical analyses let operators and regulators demonstrate to courts that they have complied with the federal "Hard Look Doctrine" in analyzing available data. Biological data is expensive in time and money to properly and usefully collect. It makes good sense to extract all the valuable information contained in these data by using robust statistical models.