Avoid, Minimize, or Mitigate: The Value of Environmental Data (Newsletter)*

September 8, 2014

Regulators require collection and submission of baseline data prior to permit issuance (e.g., NEPA documents or other operating permits), and continuing data to evaluate compliance with permit conditions. The reason is the need to determine whether the proposed project might have unacceptable environmental impacts, and whether operations have such impacts. It is common for analyses accompanying reported data to be inappropriate or superficial and not answer two critical questions. Why do observations and measurements have the values they do? Are observed variabilities natural or caused by project operations?

Answering these two questions quides operators and regulators to appropriate and effective measures that avoid, mimimize, or mitigate environmental degradation. As results, EISs are completed more quickly and permit compliance evaluations are objective and specific to projects, locations, and designated beneficial uses. Operators and regulators benefit finacially and politically.

All environmental data cannot be analyzed the same way. Chemical data are continuous with true zeros; biological data are counts or proportions. Chemical concentrations may be below laboratory method detection limits; the value is unknown. Biological data might not include a taxon previously observed at that location; it was not present or was present and not collected. These two 'zeros' are different and each must be correctly analyzed. These, and other factors, must be considered when selecting the proper statistical model for environmenal data.

The most important consideration is to fit the model to the data and not fit the data to the model. For example, time series models developed using financial and economic data assume constant measurement intervals. Environmental data almost always have irregular intervals and require time series models that allow for variable intervals.

Considerations when analyzing chemical environmental data were published in earlier newsletters. Biological environmental data analyses are more complicated. Data from aquatic and terrestrial environments vary in taxonomic identification, ease of collection or observation, and strong dependence on spatial, temporal, and geomorphic explanatory variables.

Many robust models appropriate for biological data have been developed by statisticians in recent years and are not taught in non-statistics major courses. Choice of the best model is based on the question to be answered and available data characteristics. This newsletter describes categories of the various models; future newsletters will describe them in detail.

The suite of regression models analyze cause-and-effect relationships answers questions why certain values are observed or measured. Sometimes simple linear or generalized linear regression models fits biological data, but those are exceptions rather than rules. With count data, logistic, Poisson, or quantile regression models are better fits to biological data. Generalized additive models apply to non-lnear relationships between the response and explanatory variables.

Non-numeric variables are important in explaining observed or measured environmental data. These nominal (named) or categorical variables include seasons or months, stream or drainage names, dominant plant species, and similar attributes. Nominal and categorical variables can be incorporated into mixed models such as the generalized linear mixed model (GLMM) and the generalized additive mixed model (GAMM).

Frequently, raw numbers are not as meaningful as relative proportions of biotic groups (e.g., herbivores, omnivores, carivores, and detritivores). The methods of compositional data analysis summarize statistical descriptions, compare compositional groups over space and time, and allow predictions based on explanatory variables.

It is much easier to continue using the same old methods since no one seems to care. But the investment in understanding and applying correct environmental data analyses in operational and regulatory decision-making has multiple, valuable benefits for both operators and regulators. Perhaps the most important reasons to change are that these new methods are demonstably more robust and technically sound than the 'standard' ones, and they are certainly legally defensible. Objective, quantified statistical analyses appropriately applied can shorten the permit review process, particularly with contentious and complex projects assessed under NEPA, and they can greatly reduce the likelihood of lawsuits. The last benefit is certainly of interest to politicians, legislators, senior corporate executives, financiers, and investors as well as environmental departments and regulators.

—

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