Session M8: Statistical Approaches for Assessing Water

Room C120-122

10:00 – 11:30 am

0089
M8-1

Measured Data are Uncertain: So What??

Daren Harmel

USDA-ARS, Temple, Tex., USA

The uncertainty associated with measured hydrology and water quality data is commonly acknowledged but rarely quantified in research and monitoring projects or considered in decision-making. This important aspect of water resource management and policy has been traditionally overlooked probably due to a lack of scientific understanding and the absence of a simple method to estimate uncertainty. It has also been argued that uncertainty is too difficult a concept for stakeholders and decision-makers to understand and that presentation of uncertainty (especially substantial uncertainty) will diminish usefulness of the results. These philosophical justifications for ignoring data uncertainty are, however, tenuous at best. Thus, we established a basic scientific understanding of hydrology and water quality measurement uncertainty along with a method to estimate that uncertainty. The environmental and socio-economic ramifications of decisions and policies based on these data are too great for the inherent uncertainty to continue to be ignored. We hope that our efforts have contributed to establishing a sound scientific basis for hydrology and water quality data collection that decreases resulting uncertainty and increases the efficiency of expenditures.

0006
M8-2

Multivariate Methods for Nondetects, Part 2

Dennis Helsel

Practical Stats, Castle Rock, Colo., USA

Multivariate methods such as principal component analysis, cluster analysis and multivariate tests of group differences are useful procedures when a suite of chemicals or biological community variables are analyzed for patterns. For example, a suite of trace elements might be tested to determine whether their pattern of concentrations differs among four land-use groups, or whether there is a trend in the pattern over time. A complication in the use of these methods is the presence of nondetect values, concentrations below analytical detection limits. The common solution to this issue has in the past been to substitute one-half the detection limit prior to running the procedures. This solution is actually no solution at all, introducing artificial patterns into the data that may obscure the target relationships being evaluated. The first installment of this work was presented two years ago at NWQMC 2010, highlighting possible methods rarely applied to chemical data. The ending of the story (à la Harry Potter Part 2) is presented here. Come and see who won! Not to give it away, but methods adapted from the disciplines of marine sciences and genomics are now ready to be used in statistical analysis of censored (nondetect) data. To read the story at your leisure, this material is from the author's new textbook *Statistics for Censored Environmental Data* published in January 2012 by Wiley.

0355
M8-3

Spatially Explicit Predictors of Indicators of Water Quality: Example from Wadeable Streams in the U.S.

Mostafa Shirazi and Marc Weber

US Environmental Protection Agency, Corvallis, Oreg., USA

The Environmental Protection Agency (USEPA) in collaboration with the States is assessing and reporting on the condition of surface waters in the United States using synoptic surveys and consistent field collections of water quality indicators (WQI). The survey is a probability-based design, such that inferences about all surface waters can be drawn from the sampled sites using minimal assumptions. We produced spatially explicit predictions of WQI for non-sampled locations, using the Wadeable Stream Assessment dataset (1392 WSA sites and 451 least impacted reference sites) and two sets of independent predictors: the US Natural Resources Conservation Service Dataset (STATSGO, and SSURGO) (which described the watersheds of these sites by their soil characteristics (SC)), and the US Geological Survey's National Land Cover Database (NLCD) (which further defined the land-use (LC) of the watersheds). The predictors statistically defined a non-sampled site as a nearest neighbor to one or two WSA sites. We used the mean

WQI from the WSA neighbor sites as a prediction value of one or more WQI for the non-sampled site, and a model that used the remaining WQI as co-predictors, along with the SC and LC, calculated the error of each prediction. The standard errors were measured relative to the national mean of each WQI, and they were substantially below the spatial variability (stndv/mean) of the observed data for the WSA dataset, which was <100% for six WQI and 100% to 440% for 11 of the 17 WQI, but the prediction error for the dataset was < 50% for 10 indicators, and 50% to 135% for the remaining seven WQI. The predicted WQI of 900 Level IV Ecoregions was compared with reference sites, which displayed regions of potential water quality and water quality vulnerability in the conterminous U.S.

0443
M8-4

Evaluating Wetland Health: Avoiding Indexes via a Multivariate Latent Variable Model

Jennifer Hoeting and Erin Schliep

Colorado State Univ., Fort Collins, Colo., USA

We examine the problem of spatially correlated, multivariate ordinal data in the assessment of health of an ecological system. A common approach is to use an index to combine various biological or physiochemical metrics that are typically ordinal-scaled variables. For example, stream health is often measured by the Index of Biotic Integrity (IBI), which is the weighted-sum of a collection of metrics at a site. Indices such as these are criticized because the weights assigned to each metric are determined arbitrarily. In this work, we overcome this shortcoming. We propose a latent model for multivariate ordinal data and consider several extensions. We estimate parameters using Bayesian methods. We demonstrate our methods using a US Environmental Protection Agency supported study of wetland health in Colorado. The model allows us to rank wetland health at multiple locations, predict health at locations not visited in the field sample, and rank the importance of a suite of potential threats and stressors.