Geophysical Research Abstracts
Vol. 13, EGU2011-6577, 2011
EGU General Assembly 2011

Generalised linear mixed models for predicting the probability of exceeding water quality guidelines

Jason Lessels, Michael Nelson, Thomas Bishop, and Rutger Vervoort
Faculty of Agriculture, Food and Natural Resources, University of Sydney, Australia

Water quality guidelines are used worldwide to assess the health of catchments yet surprisingly little research effort has been focused on examining methods for estimating the number of times in a time period that the guidelines are exceeded. This is in contrast to load estimation methods which have received far more attention. In Australia a typical approach by catchment authorities is to count the number of observations where the guidelines have been exceeded and use this to give a percentage of samples above the threshold. This approach is highly flawed as it is based on a small number of samples which do not necessarily reflect the entire range of flow conditions. This is especially problematic in Australia where flow and water quality are heavily correlated, and flow regimes are highly variable. In addition this approach does not provide an estimate of the probability of the guidelines for particular times, such as unsampled flow events.

In this research we describe the use of generalised linear mixed models (GLMM) to estimate the probability of exceeding water guidelines. To illustrate the approach we use a dataset from the Wollondilly river which provides a significant portion of the drinking water for Sydney, Australia. At this site a flow gauge records flow at hourly intervals and the total phosphorous has been measured irregularly using both regular monthly sampling, since 1991, and automatic sampling based on changes in flow, since 2001.

GLMMs offer many advantages, firstly they are statistically correct in that a binomial distribution is used to model the response and the predictions are of the probability of the guidelines being exceeded. Furthermore, the method does not assume independent errors and accounts for auto-correlation through time between water quality samples. If independent errors were assumed the prediction variances would be underestimated. Secondly it takes advantage of the intensively measured flow data to predict the exceedance probability at the same hourly measurement interval and it also uses adjacent water quality observations in time, to improve the predictions through interpolation. Finally GLMMs provide a confidence interval around the predicted probability which is a measure of the reliability of the predictions. The provision of a confidence intervals is key to using GLMMs to improve monitoring schemes by optimising when samples should be taken.

Aside from using GLMMs to guide improvements in the design of water quality monitoring programs, future work is needed on:
-testing different statistical methods for obtaining confidence intervals;
-provide guidelines for catchment managers on how to interpret the predicted probabilities, and their associated confidence intervals, to guide management decisions.