Utah's Biological Assessment Model for Streams

Introduction

Utah's beneficial uses for aquatic life require the protection of fish (cold water or warm water species) and the organisms on which they depend (UAC R317-2-6.3). DWQ historically assessed these beneficial uses using water chemistry sampling and associated standards that are protective of aquatic organisms. DWQ now uses an empirically based model that directly assesses support of aquatic life uses by quantifying the integrity of macroinvertebrate assemblages. Measuring biological communities directly has the advantage of integrating the combined effects of all pollutants, which allows a direct examination of how pollutants are interacting to affect the condition of a stream ecosystem (Karr,1981). Moreover, because aquatic macroinvertebrates spend most of their life in aqueous environments, they are capable of integrating the effects of stressors over time, providing a measure of past and transient conditions (Karr and Dudley,1981).

Biological assessments are often conducted by comparing the biological assemblage observed at a site with the expected biological assemblage. Ideally, these comparisons are made using historical data to measure changes to the current biological community. However, in most cases, historical data are not available. As a result, biological conditions representing least-human-caused disturbance are typically set using reference sites as benchmarks or controls. The biological integrity of sites can then be evaluated by comparing the biological composition observed at a site against a subset of ecologically similar reference sites. Such comparisons are collectively referred to as biological assessments.

Reference sites in aquatic biological assessments are selected to represent the best available condition for waterbodies with similar ecological, physical, and geographical characteristics (Hughes et al.,1986; Suplee et al.,2005). Conditions at reference sites selected for water quality programs vary regionally depending on adjacent historical land use. For example, reference sites in Utah mountains are generally more pristine than in valleys. As a result, there are more biological benchmarks in areas of the state that receive less-human-made disturbance than those with more disturbances.

A numeric index is a useful tool that quantifies the biological integrity, or biological beneficial use, of stream and river segments. Data obtained from biological collections are complex, with hundreds of species that vary spatially and temporally found throughout Utah. Similarly, the physical template on which biota depends also varies considerably across streams. A robust index of biological integrity should simultaneously account for naturally occurring physical and biological variability and summarize these conditions through a single, easily interpretable number (Hawkins, 2006; Hawkins et al. 2010).

River Invertebrate Prediction and Classification System Models

DWQ uses the River Invertebrate Prediction and Classification System (RIVPACS) model approach to quantify biological integrity (Wright, 1995). RIVPACS is a classification of freshwater sites based on macroinvertebrate fauna. It was first derived in 1977 and has subsequently been used in numerous biological assessment programs worldwide. In the early

1970s, scientists and water managers recognized a need to understand the links between the ecology of running waters and macroinvertebrate communities. A four-year project was initiated to create a biological classification of unpolluted running waters in Great Britain based on the macroinvertebrate fauna (Clarke et al.,1996; Furse et al.,1984; Moss et al.,1999; Wright,1995).

Over the past 30 years, equivalent RIVPACS models have been developed for aquatic ecosystems throughout the world, including Australia (Davies et al., 2000; Marchant and Hehir, 2002; Metzeling et al., 2002) and Indonesia (Sudaryanti et al., 2001). Additionally, scientists in the United States have developed RIVPACS models to assess the biological integrity of the country's aquatic habitats (Hawkins et al., 2000; Hawkins and Carlisle, 2001).

RIVPACS models compare the list of taxa that are observed (O) at a site to the list of taxa expected (E) with the least-human-caused disturbance for a similar site to quantify biological condition. Predictions of E are obtained empirically from reference sites that together are assumed to encompass the range of ecological variability observed among streams in the region where the model was developed. In practice, these data are expressed as the ratio O/E, the index of biological integrity.

Desert Site
O = 7
E = 10
0.70

Mountain Site
O = 21
E = 30
0.70
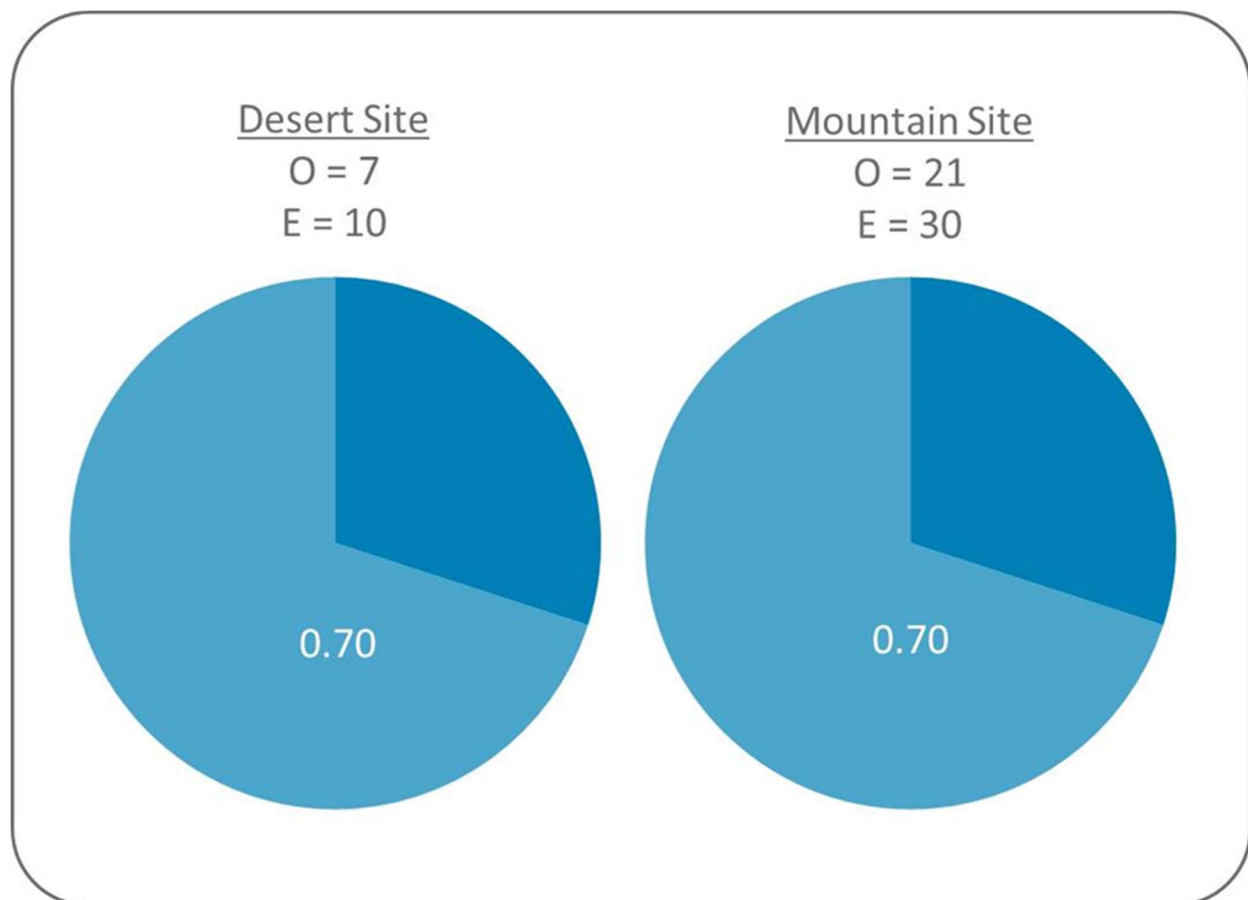
O/E has some very useful properties as an index of biological condition. First, it has an intuitive biological meaning. Species diversity is considered the ecological capital on which ecosystem processes depend; therefore, O/E can be easily interpreted by researchers, managers, policy-makers, and the public. Second, O/E is universally spatial, which allows direct and meaningful comparison throughout the state on a site-specific scale. This is particularly important for Utah,

where streams vary considerably from high-altitude mountain environments to arid desert regions. Third, its derivation and interpretation do not require knowledge of stressors in the region as it is simply a biological measuring tool. Finally, the value of O/E provides a quantitative measure of biological condition.

Model Construction and Performance

Construction of a RIVPACS model for Utah began in 2002 and involved the development and evaluation of dozens of models. Details of model development procedures can be found elsewhere (Clarke et al.,1996; Moss et al.,1999; Wright et al.,1993; Wright1995). Additionally, specific detailed instructions can be viewed at the Western Center for Monitoring and Assessment of Freshwater Ecosystems website and there are numerous resources found online. DWQ's model was verified and reconstructed by the USU BugLab that can provide O/E output for samples if appropriate field and lab procedures were followed. A brief summary is provided here to help the reader better understand Utah's model results and subsequent assessments.

Predictions of expected "E" taxa are obtained empirically from reference site collections made throughout Utah. Reference sites represent the reference conditions in different biogeographical settings throughout the state. The initial list of candidate reference sites is independently ranked by different scientists familiar with the waterbodies. Only reference sites with a consensus representing best available conditions are used in model development. Subsequent reference sites are added using scores from reference-scoring metrics developed during site visits and averaged with independent rankings from field scientists.

Some of the calculations used to obtain the list of expected taxa are complex. A heuristic description of the steps involved in predicting "E" provides some context of the assessment methods. The first step in model development is to classify reference sites into groups of sites with similar taxonomic composition using a cluster analysis. Next, models are developed based on watershed descriptors such as climatic setting, soil characteristics, and stream size to generate equations that predict the probability of a new site falling within each group of reference sites. These equations account for environmental heterogeneity and ensure that when a new site is assessed, it is compared against ecologically similar reference sites. When a new site is assessed, predictions of group membership are then coupled to the distributions of taxa across groups of reference sites to estimate the probability of capturing (Pc) each taxon from the regional pool of all taxa found across all reference sites. E is then calculated as the sum of all taxa Pcs that had a greater than 50% chance of occurring at a site given the site's specific environmental characteristics. Using a Pc limit set at greater than 50% typically results in models that are more sensitive and precise, which results in a better ability to detect biological stress (Hawkins et al., 2000; Simpson and Norris, 2000; Ostermiller and Hawkins, 2004; Hawkins, 2006; Van Sickle et al., 2007, Hawkins et al., 2015; Hawkins and Yuan, 2016; Mazor et al., 2016).

The accuracy and precision of RIVPACS models depend in part on the ability of the models to discriminate among groups of biologically similar reference sites. An extensive list of 74 GIS-based watershed descriptors is evaluated for potential predictor variables in models that predict the probability of membership within biological groups for sites not used in model construction. Site-specific, GIS-based predictor variables, such as soils, meteorology, and geography, instead of field-derived descriptors, are evaluated for a couple of reasons. First, GIS-based descriptors are unlikely to be influenced by human disturbance and are therefore unlikely to bias estimates of expected conditions (Hawkins, 2004). Second, these predictors are easily obtained for any

location on a site-specific basis. This allows inclusion of additional macroinvertebrate samples collected by others. Various subsets of potential predictors are evaluated in an iterative, analytical process that explores different combinations of predictors to explain the biological variability among reference sites. The current RIVPACS model used by DWQ includes 15 variables that resulted in the most precisely predictive model (see table).

General Category	Description
Geography	Mean watershed elevation (meters) from National Elevation Dataset
Geography	Minimum watershed elevation (meters) from National Elevation Dataset
Geography	Watershed area in square kilometers
Geography	Latitude of the sample location.
Climate	Watershed average of the mean day of year (1–365) of the first freeze derived from the PRISM data
Climate	Watershed average of the annual mean of the predicted mean monthly precipitation (millimeters) derived from the PRISM data
Climate	Watershed average of the annual maximum of the predicted mean monthly precipitation (millimeters) derived from the PRISM data
Climate	Watershed average of the annual mean of the predicted mean monthly air temperature derived from PRISM data

Climate	Average of the annual mean of the predicted maximum monthly air temperature at the sample location derived from PRISM data
Climate	Watershed average of the annual mean of the predicted maximum monthly air temperature derived from PRISM data
Climate	Watershed average of the annual mean of the predicted minimum monthly air temperature derived from PRISM data
Climate	Watershed average of the annual mean of the predicted mean monthly relative humidity derived from PRISM data
Climate	Average of the annual mean of the predicted mean monthly air temperature at the sample location derived from PRISM data
Climate	Watershed maximum of mean 1961–1990 annual number of wet days
Vegetation	Watershed maximum of mean 2000–2009 annual enhanced vegetation index

The RIVPACS model used for the current assessments was reconstructed to accommodate broader spatial and temporal data. Models used earlier were limited to samples from streams ranging from second to fifth order and were collected during a fall window of September–November. The updated model accepts data collected from first- to eighth-plus-order rivers and streams with no limitations on season of collection. In addition, new predictor variables were tested, and new and updated reference site data were included. However, the taxon levels required adjustment to include data collected from agencies using different taxonomic laboratories. This resulted in a coarser resolution of taxonomy. However, the resulting model was capable of scoring nearly 1,800 samples collected across the state by various agencies.

The updated model is nearly as accurate and precise as previous models. If the model was perfectly accurate and precise, the O/E score for all reference sites would equal 1.00. Instead, reference O/E values are typically spread in a roughly normal distribution centered on 1.00 (Wright, 1995). Model precision is often expressed as the standard deviation (SD) of reference

O/E values, with lower SDs indicating higher model precision. The RIVPACS model used for the current IR assessments has an SD of 0.19, which is within the range of water quality models used for biological assessment purposes. The precision was likely affected by the coarser resolution of taxonomy and the inclusion of a few large river sites as reference. The average reference O/E score for the current model is 1.00, which means that the model has high precision calculating O/E values. The accuracy of the model was evaluated by examining the distribution of reference O/E scores across environmental settings and determining that reference O/E values are not biased by stream size, elevation, or ecoregion because the model performed similarly under a broad range of physical conditions.

References

Clarke, R.T., M.T. Furse, J.F. Wright, and D. Moss. 1996. Derivation of a biological quality index for river sites: comparison of the observed with the expected fauna. Journal of Applied Statistics 23:311–332.

Davies, N.M, R.H. Norris, and M.C. Thoms. 2000. Prediction and assessment of local stream habitat features using large-scale catchment characteristics. Freshwater Biology 45:343–369.

Furse, M.T., D. Moss, J.F. Wright, and P.D. Armitage. 1984. The influence of seasonal and taxonomic factors on the ordination and classification of running-water sites in Great Britain and on the prediction of their macro-invertebrate communities. Freshwater Biology 14:257–280.

Hawkins, C.P., Cao, Y. and Roper, B., 2010. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. Freshwater Biology, 55(5), pp.1066-1085.

Hawkins, C.P., Mykrä, H., Oksanen, J. and Vander Laan, J.J., 2015. Environmental disturbance can increase beta diversity of stream macroinvertebrate assemblages. Global Ecology and Biogeography, 24(4), pp.483-494.

Hawkins, C.P, R.H. Norris, J.N. Hogue, and J.W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications 10:1456–1477.

Hawkins, C.P., Olson, J.R. and Hill, R.A., 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. Journal of the North American Benthological Society, 29(1), pp.312-343.

Hawkins, C.P. and Yuan, L.L., 2016. Multitaxon distribution models reveal severe alteration in the regional biodiversity of freshwater invertebrates. Freshwater Science, 35(4), pp.1365-1376.

Hughes, R.M., D.P. Larsen, and J.M. Omernik. 1986. Regional reference sites: a method for assessing stream potential. Environmental Management 5:629–635.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6:21–27.

Karr, J.R., and D.R. Dudley. 1981. Ecological perspectives on water quality goals. Environmental Management 5(1):55–68.

Moss, D. J.F. Wright, M.T. Furse, and R.T. Clarke. 1999. A comparison of alternative techniques for prediction of the fauna of running-water sites in Great Britain. Freshwater Biology 41:167–181.

Marchant, R., and G. Hehir. 2002. The use of AUSRIVAS predictive models to assess the

response of lotic macroinvertebrates to dams in south-east Australia. Freshwater Biology 43:1022–1050.

Mazor, R.D., Rehn, A.C., Ode, P.R., Engeln, M., Schiff, K.C., Stein, E.D., Gillett, D.J., Herbst, D.B. and Hawkins, C.P., 2016. Bioassessment in complex environments: designing an index for consistent meaning in different settings. Freshwater Science, 35(1), pp.249-271.

Metzeling, L., D. Robinson, S. Perris, and R. Marchant. 2002. Temporal persistence of benthic invertebrate communities in south-eastern Australian streams: taxonomic resolution and implications for the use of predictive models. Marine and Freshwater Research 53:1223–1234.

Ostermiller, J.D. and Hawkins, C.P., 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of the North American Benthological Society, 23(2), pp.363-382.

Simpson, J.C., and R.H. Norris. 2000. Biological assessment of river quality: development of AusRivAS models and outputs. In Assessing the Biological Quality of Fresh Waters, edited by J.F. Wright, D.W. Sutcliffe, and M.T. Furse, pp. 125–142. Ambleside, United Kingdom: Freshwater Biological Association.

Sudaryanti, S., Y. Trihadiningrum, B.T. Hart, P.E. Davies, C. Humphrey, R.H. Norris, J. Simpson, and L. Thurtell. 2001. Assessment of the biological health of the Brantas River, East Java, Indonesia using the Australian River Assessment System (AUSRIVAS) methodology. Aquatic Ecology 35(2):135–146.

Suplee, M., R. Sada de Suplee, D. Feldman, and T. Laidlaw. 2005. Identification and Assessment of Montana Reference Streams: A Follow-Up and Expansion of the 1992 Benchmark Biology Study. Helena, Montana: Montana Department of Environmental Quality.

Wright, J.F. 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. Australian Journal of Ecology 20:181–197.

Van Sickle, J., Larsen, D.P. and Hawkins, C.P., 2007. Exclusion of rare taxa affects performance of the O/E index in bioassessments. Journal of the North American Benthological Society, 26(2), pp.319-331.