Comparison of the performance of different bioassessment methods: similar evaluations of biotic integrity from separate programs and procedures

David B. Herbst[1]

Sierra Nevada Aquatic Research Laboratory, University of California, Route 1, Box 198, Mammoth Lakes, California 93546 USA

Erik L. Silldorff[2]

Princeton Hydro, 1108 Old York Road, Suite 1, P.O. Box 720, Ringoes, New Jersey 08551 USA

Abstract. Regional bioassessment programs of states, various federal agencies, and other governmental and private groups often use different methods to collect and analyze stream invertebrate samples. This lack of uniformity has created concern and confusion over the comparability of disparate sources of data, but few studies have attempted to evaluate differences in performance between methods or to reconcile the results produced from different methods. We conducted concurrent sampling at 40 sites in the eastern Sierra Nevada of California using 3 bioassessment methods to obtain directly comparable data sets. The riffle-based methods (University of California Sierra Nevada Aquatic Research Laboratory [UC-SNARL, Lahontan Water Board], California Stream Bioassessment Protocol, and US Forest Service Region 5) differed at each stage from field sample collection to laboratory processing and data analysis. We used a performance-based methods system to compare precision, uniformity, discrimination, accuracy, and correlations among multimetric Index of Biotic Integrity (IBI) scores and multivariate River Invertebrate Prediction and Classification System (RIVPACS)-type observed/expected (O/E) ratios. Reference and test sites were identified using local and upstream-watershed disturbance criteria, and invertebrate community measures and models were then developed to discriminate between reference and test sites. The more-intensive UC-SNARL method showed slightly, but consistently, greater sensitivity for discriminating impairment than the other 2 methods. The UC-SNARL method produced greater differences between reference- and test-site means relative to lower reference-site standard deviations than the other 2 methods. However, assessment scores were highly correlated among methods and distinguished reference from test sites with similar accuracy among methods despite the slight differences in performance. Our results show that differing bioassessment methods can yield very similar, effective discrimination of impaired biological condition even though they have multiple differences in field and laboratory protocols (mesh size, replication, area sampled, taxonomic resolution, total counts). Moreover, this conclusion did not depend on the approach taken to data analysis because both multimetric IBIs and multivariate RIVPACS-type O/Es were in close agreement. Methodological uniformity is important when coordinating monitoring programs, but our results suggest that data from multiple sources could potentially be used interchangeably and for cross-validation of assessments of stream biological integrity.

Key words: bioassessment, impairment detection, methods comparison, metric precision, multimetric IBI, RIVPACS, Sierra Nevada, stream macroinvertebrates.

Surveys of the different stream bioassessment protocols used among federal, state and local programs show considerable variation in the procedures and tools used to collect and process samples (Gurtz and Muir 1994, Carter and Resh 2001). Comparisons of

the data derived from collections taken with various types of sampling equipment, subsampling counts, and levels of taxonomic resolution have provided a basis for evaluating some of the field and laboratory methods in use (Resh and McElravy 1993, Resh and Jackson 1993, Barbour and Gerritsen 1996, Courtemanch 1996, Vinson and Hawkins 1996, Lenat and Resh 2001). The techniques used to analyze bioassess-

[1] E-mail addresses: herbst@lifesci.ucsb.edu
[2] esilldorff@princetonhydro.com