Ecological Indicators 36 (2014) 470–482

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Enhancing the performance and interpretation of freshwater biological indices: An application in arid zone streams

Jacob J. Vander Laan*, Charles P. Hawkins

Department of Watershed Sciences, Western Center for Monitoring and Assessment of Freshwater Ecosystems, and Ecology Center, Utah State University, Logan, UT 84322-5210, USA

ARTICLE INFO

Article history:
Received 13 February 2013
Received in revised form 30 August 2013
Accepted 3 September 2013

Keywords:
Arid environments
Bioassessment
Biological indices
Community assembly
Environmental monitoring
Habitat isolation
Meta-communities
Models
Incomplete sampling
Evenness
Stream ecosystems

ABSTRACT

We assessed the performance of biological indices developed for invertebrate assemblages occurring in arid zone streams: a multimetric index (MMI) and an O/E index of taxonomic completeness. Our overall goal was to advance our understanding of the factors that affect performance and interpretation of biological indices. Our specific objectives were to (1) develop biological indices that are insensitive to natural environmental gradients, (2) develop a general method to determine if the biological potential of an assessed site is adequately represented by the population of reference sites, (3) develop a robust method to select metrics for inclusion in MMIs that ensures maximum independence of metrics, and (4) determine if a fundamental sample property (the evenness of taxa counts within a sample) affects index performance. Random Forest modeling revealed that both individual metrics and taxa composition were strongly associated with natural environmental heterogeneity, which meant both the MMI and O/E index needed to be based on site-specific expectations. We produced a precise, responsive, and ecologically robust MMI by using principal components analysis to identify 7 statistically independent metrics from a list of 31 candidate assemblage-level metrics. However, the O/E index we developed was relatively imprecise compared with O/E indices developed for other regions. This imprecision may be the consequence of low predictability in local taxa composition associated with the relatively high spatial isolation of aquatic habitats within arid regions. We were also able to assess the likelihood that the biological potential of assessed sites were adequately characterized by the population of reference sites by developing and applying a multivariate, nearest-neighbor test that determined if an assessed site occurred within the environmental space of the reference site network. This approach is robust and applicable to all biological indices. We also demonstrate that the evenness of taxa counts within a sample is positively related to estimates of sample taxa richness and thus the scores of both indices. The relationship between richness and sample evenness can potentially compromise inferences regarding biological condition, and post hoc adjustments for the effects of evenness on index scores might be desirable. Further improvements in the performance and interpretation of biological indices will require simultaneous consideration of the effects of incomplete sampling on characterization of biological assemblages and the physical and biological factors that influence community assembly.

1. Introduction

The technical foundations on which freshwater ecological assessments depend have substantially improved over the last 30 years. Especially significant advances have been made in defining, characterizing, applying, and interpreting both the benchmarks (or reference state) used to determine ecological condition (e.g., Stoddard et al., 2006; Hawkins et al., 2010a,b) and the ecological endpoints used to measure condition (Barbour et al., 1999; Karr, 1999; Hawkins et al., 2000b; Ruaro and Gubiani, 2013). However, challenges still remain regarding how to best translate general concepts to application, especially when developing and applying scoring tools for use under specific environmental settings. These challenges include (1) improving accuracy and precision when estimating biological reference conditions, (2) ensuring that we do not inappropriately assess individual locations whose natural environmental settings are outside of the range of naturally occurring conditions represented by the network of reference sites, (3) improving the behavior and robustness of the biological indices used to assess ecological endpoints of interest by ensuring that the components included in these indices represent meaningful aspects of biological integrity, and (4) improving our understanding of how sample properties can affect interpretation of index scores.

* Corresponding author. Tel.: +1 6162600040.
E-mail addresses: jacob.vl@usu.edu (J.J. Vander Laan), chuck.hawkins@usu.edu (C.P. Hawkins).

1470-160X/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ecolind.2013.09.006

Addressing these challenges is especially important in understudied regions and environments. In this paper, we describe research based on data collected from streams in Nevada (NV), the most arid state in the USA. Our general goal for this study was to advance our understanding of the factors that affect performance and interpretation of biological indices. In particular, we focused on four objectives: (1) develop ecologically and statistically robust biological indices that are insensitive to natural environmental gradients, (2) develop a general method to determine if the biological potential of an assessed site is adequately represented by the population of reference sites, (3) develop a robust method to select metrics for inclusion in MMIs that ensures maximum independence of metrics, and (4) determine if the evenness of taxa counts within a sample affects estimates of sample richness and index scores.

One of the biggest challenges for any bioassessment is accounting for naturally occurring spatial and temporal variation in assemblage composition (Karr and Chu, 1997; Barbour et al., 1999; Hawkins et al., 2010a,b). Failure to account for the effects of natural variability on the abundance and distribution of biota can lead to confounding of natural and anthropogenic effects on assemblages. This confounding can lead to biased assessments and result in increased type I and type II errors of inference, i.e., a site being assessed as impaired when it is not and vice versa. O/E indices and MMIs generally differ in how they account for natural variability in assemblages and define biological expectations. Biological expectations for O/E indices are set by modeling natural variability among assemblages and predicting site-specific expectations of the probability of observing individual taxa. MMIs have traditionally relied on some type of regionalization to account for the natural variation of metric values among reference sites and thus set expectations for biological condition at assessed sites (Barbour et al., 1999). However, these regionalizations often account for little of the biological variation that occurs between sites (Hawkins et al., 2000a, 2010a). Recently, some MMI developers have also used modeling techniques to adjust metric expectations for natural gradients (e.g. Baker et al., 2005; Pont et al., 2006; Cao et al., 2007; Hawkins et al., 2010a), which can increase MMI accuracy and precision (e.g. Cao et al., 2007; Hawkins et al., 2010a).

Accurate bioassessments require a comparison of the biota observed at a site with an estimate of that site's biological potential. Sites with similar natural environmental settings are considered to have similar potential. Accurate determination of biological expectations requires that the biological potential of any assessed site is adequately represented within the pool of reference sites. For O/E indices, Moss et al. (1987) developed a statistical test to determine if the combination of values for those specific environmental attributes used to predict reference condition biota (E) measured at an assessed site occur within the pool of reference sites. However, their method, which is based on the discriminant functions used to predict *E* in the RIVPACS model, is not applicable to other modeling approaches or types of indices. It would be useful to have a general approach to identifying sites that lack appropriate reference sites that could be used with all indices. MMIs have generally focused on ensuring that some minimum set of reference sites occurs within all regions or typologies, but how well those reference sites represent environmental gradients within regions has typically not been considered. A more broadly applicable method for identifying sites whose environmental conditions are not represented within the set of reference sites would help ensure the accuracy of a variety of bioassessment indices when applied to sites of unknown condition.

Stream bioassessments typically use one of two types of biological indices: multimetric indices (MMIs) and observed to expected (O/E) taxa ratios (Cao and Hawkins, 2011). An MMI aggregates several measures of invertebrate assemblage attributes (Karr and Chu, 1997). The types of attributes used in MMIs often include measures of taxonomic richness, biological diversity and assemblage composition (Stoddard et al., 2008). Individual metrics that differentiate between reference and degraded condition are standardized and aggregated into a single measure of biological condition. O/E indices assess the taxonomic completeness of a site by comparing observed and expected taxa lists (e.g. Wright, 1995; Hawkins, 2006). The taxa expected at a site are predicted by identifying relationships between taxonomic composition and environmental gradients at pre-defined reference sites, and biological condition at assessed sites is measured as the number of expected taxa (E) that are observed (O) at a site, usually expressed as a ratio, O/E (Wright, 1995). Benthic macroinvertebrate (BMI) based bioassessment indices have become increasingly important and widely used tools for assessing the biological condition of freshwater ecosystems, but index performance varies widely, and the sources of this variation are not well understood.

Metric redundancy is generally thought to compromise MMI performance (Van Sickle, 2010), and index developers have tried to minimize redundancy in several ways. Approaches for identifying redundant metrics include setting a maximum value for the allowed correlation between metrics in an MMI (Barbour et al., 1999; Hering et al., 2006; Stoddard et al., 2008), using best professional judgment to classify metrics into ecologically independent categories (Hering et al., 2006; Stoddard et al., 2008), and objectively clustering metrics based on their correlations (Cao et al., 2007). Typically in these approaches, a single or equal number of metrics from each set of redundant metrics is selected for the final index. However, these approaches do not necessarily ensure statistical or biological independence of metrics and require the application of an arbitrary correlation coefficient cutoff. Van Sickle (2010) demonstrated that the maximum correlation between two metrics in an MMI has little or no effect on index performance and instead suggests that minimizing the mean correlation between metrics is a more effective method of increasing index performance. An objective method for reducing redundancy among metrics could help ensure the development of ecologically and statistically robust MMIs.

Both O/E indices and MMIs are based in whole, or in part, on measures of taxonomic richness. Therefore, any sample property, such as sample evenness, that affects estimates of richness may affect index performance. For example, the taxonomic richness of stream invertebrate samples typically increases asymptotically with the number of individuals in a sample (Vinson and Hawkins, 1996). However, this asymptote will be approached more rapidly in highly even assemblages than in less even ones (Gotelli and Colwell, 2001). Therefore, differences in sample evenness may lead to differences in estimates of sample richness and composition. This dependency of estimated richness on evenness will be particularly problematic for the small, fixed-count samples commonly used in bioassessments (Cao and Hawkins, 2005). Differences in sample evenness could therefore increase bias and decrease precision of O/E indices. Because MMIs are often comprised of richness attributes, MMI scores may also be affected by sample evenness.

2. Methods

2.1. Study area and data

We used a NV (USA) Department of Environmental Protection (NDEP) dataset consisting of macroinvertebrate samples collected at 500 stream sites throughout NV and nearby surrounding areas to develop MMI and O/E indices for the state (Fig. 1). Data included samples from 415 sites of unknown biological condition collected by the NDEP and samples from 85 sites previously determined to be in reference condition that were sampled by either the Western Center for Monitoring and Assessment of Freshwater Ecosystems