Ecological Indicators 71 (2016) 533–543

Contents lists available at ScienceDirect

Ecological Indicators

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

Original Articles

Modeling wetland plant metrics to improve the performance of vegetation-based indices of biological integrity

Robin C. Jones[a], Charles P. Hawkins[a,*], M. Siobhan Fennessy[b], Jacob J. Vander Laan[a]

[a] Department of Watershed Sciences, Western Center for Monitoring and Assessment of Freshwater Ecosystems, Ecology Center, Utah State University, Logan, UT 84322-5210, United States
[b] Department of Biology, Kenyon College, Gambier, OH 43022, United States

ARTICLE INFO

Article history:
Received 7 May 2015
Received in revised form 19 July 2016
Accepted 20 July 2016

Keywords:
Reference condition
Index of biological integrity
Predictive models
Random forests
Vegetation
Wetlands

ABSTRACT

The objective of this study was to determine if the accuracy and precision of wetland plant indices of biological integrity (IBIs) could be improved through the use of modeling techniques. To do this, we developed a modeled vegetation IBI (MVIBI) based on metrics previously used to develop vegetation indices of biological integrity (VIBIs) for Ohio wetlands (e.g. % invasive grass, % sensitive species, shrub richness). We selected 82 emergent, forested, and shrub-dominated reference sites distributed across the State of Ohio and built Random Forest models to predict plant metric scores at reference wetlands from naturally occurring environmental features related to climate, hydrology, geology, soils, and landscape position. The models explained between 14 and 52% of the variance in the scores of 21 metrics indicating that variation in wetland plant assemblages was significantly associated with naturally occurring environmental gradients. We used principal component analysis to identify ten groups of statistically independent metrics and selected one metric from each group that discriminated most strongly between reference and most degraded sites based on t-scores. Two axes did not contain discriminating metrics so we used eight metrics in the MVIBI. Analysis of variance of reference site MVIBI scores indicated that we could use one distribution of reference site scores to assess multiple wetland types, thus eliminating the need to separately designate wetland types. We used the MVIBI to assess 170 test sites and compared the accuracy, precision, responsiveness, and sensitivity of the MVIBI to those of the original VIBIs. The MVIBI was up to twice as accurate and precise as the original VIBIs, indicating that modeling can be used to improve the performance of vegetation-based IBIs. The use of model-based IBIs for wetland plants should reduce assessment errors associated with natural variation in plant metrics and should increase confidence in wetland assessments.

1. Introduction

Accurate biological assessments depend on correctly characterizing the reference condition, of assessed sites (Bailey et al., 2004; Hawkins et al., 2010b; Herlihy et al., 2008; Hughes et al., 1986; Stoddard et al., 2006). The reference condition is an estimate of the biological potential, i.e. the best biological condition, that a site can achieve. Accurately and precisely estimating the reference condition of individual sites is a significant challenge, and the approach used to characterize the reference condition can influence the reliability of assessments (Bailey et al., 2004; Hawkins et al., 2010a; Stoddard et al., 2008). The biological reference condition theoretically represents the range of natural biological variation that a specific site would exhibit over time (Hawkins et al., 2010b; Stoddard et al., 2006), but natural resource managers are rarely able to monitor individual sites over several years to estimate this range. Instead, the reference condition is usually estimated via a space-for-time substitution where the biological variation among many reference sites (i.e. sites in the least disturbed or best available condition) is used to estimate the range of natural variation expected at individual test sites. However, the range of natural environmental variation among reference sites (i.e. that associated with natural gradients in climate, geology, soils, etc.) can be large, and if index developers fail to control for naturally occurring spatial variation, assessments can suffer from either unacceptably high type I (false positive) or type II (false negative) errors of inference.

Most plant-based indices of biological integrity (IBIs) used to assess wetland condition are based on reference condition estimates in which some form of classification is used to control for natural variation in biological attributes (e.g. DeKeyser et al., 2003;

* Corresponding author.
E-mail address: chuck.hawkins@usu.edu (C.P. Hawkins).

http://dx.doi.org/10.1016/j.ecolind.2016.07.030

Miller et al., 2006; Reiss, 2006; Rothrock et al., 2008). Classification schemes are typically geographical, typological, or both, and are used to identify groups of sites that are assumed to be environmentally and biologically similar (Bailey et al., 2004; Brinson, 1993; Karr and Chu, 1999). For example, Miller et al. (2006) developed an IBI specifically for headwater wetlands in the Ridge and Valley Province of Pennsylvania. The accuracy of this approach is dependent on how well the classification meets the following assumptions: (1) it minimizes naturally occurring variation in plant attributes within classes, (2) the range of biological conditions observed at sites within a class is representative of the temporal variation typically observed at individual sites within a class, and (3) index scores at assessed sites can potentially fall outside of the range of scores found at reference sites. However, many wetland IBIs may not meet these assumptions as geographical and typological classifications sometimes fail to account for a large portion of the natural biological variation among sites (Hawkins et al., 2000; Hawkins and Vinson, 2000; Heino et al., 2002). Another consequence of using classification is that indices may not be readily transferrable to other regions or types of wetlands.

The use of wetland type-specific IBIs, in which classes are defined based on dominant vegetation, can also be problematic for other reasons. For example, many wetlands contain several different types of plant assemblages (i.e., mixed assemblages) and their composition can be dynamic over time. When assemblages are mixed, wetland type determinations can often be subjective (Cronk and Fennessy, 2001; Stapanian et al., 2013) and prone to error. Additionally, wetland type classifications can change as vegetation assemblages shift in response to changes in both the abiotic and biotic environment (Cronk and Fennessy, 2001; Magee and Kentula, 2005; Miller and Zedler, 2003). If reference-quality wetlands within a wetland type contain mixed plant assemblages, or are in transition from one wetland type to another, then estimates of the reference condition based on wetland classification will likely be imprecise, and assessments may be unable to distinguish anthropogenic alteration from natural variation.

Modeling is an alternative approach for estimating the range of natural biological variation expected at individual sites under reference conditions. This approach employs statistical models that relate species composition or biological attributes (metrics) at reference sites to naturally occurring environmental attributes. Such models can then be used to predict species or assemblage traits that should occur at individual sites in the absence of anthropogenic stress. The values of biological metrics observed at individual test sites are then adjusted (subtracted from predicted values) to account for the effects of natural environmental gradients. Models have been widely used to predict reference conditions for lake, stream, and river ecosystems (e.g. Cao et al., 2007; Moss et al., 1987; Hawkins et al., 2010a,b; Stevenson et al., 2013) and can identify metrics that are broadly applicable across large regions and many types of sites (e.g., Stoddard et al., 2008). Model-adjusted indices can be developed at large spatial scales because models can account for more biological variation associated with naturally occurring environmental gradients than classifications (Hawkins et al., 2000; Hawkins and Vinson, 2000; Paulsen et al., 2008; Stoddard et al., 2008). Use of a common set of metrics can also improve bioassessment comparability (Cao and Hawkins, 2011). However, only a few studies describe model-based approaches used for either wetland systems or vascular plant assemblages (Aguiar et al., 2011; Cohen et al., 2005; Sifneos et al., 2010).

Index performance (e.g., accuracy and precision) is critical if biological assessments are to provide credible evaluations of biological condition. Current plant-based IBIs may not fully meet the performance needs of wetland managers because of potential inaccuracies, limited applicability, and lack of comparability associated with the use of class-specific metrics and indices. The goal of this study was to determine if current, plant-based wetland IBIs can be improved by modeling relationships between plant metrics and environmental features known to influence wetland plant assemblages such as landscape position, hydrology, and soil type (Cronk and Fennessey, 2001; Tiner, 1998). Our specific objectives were to determine if modeling could: (1) improve index precision, (2) reduce local index bias associated with natural environmental differences among sites, (3) improve index applicability by incorporating multiple wetland types into one index, and (4) improve comparability of assessments across different types of wetlands.

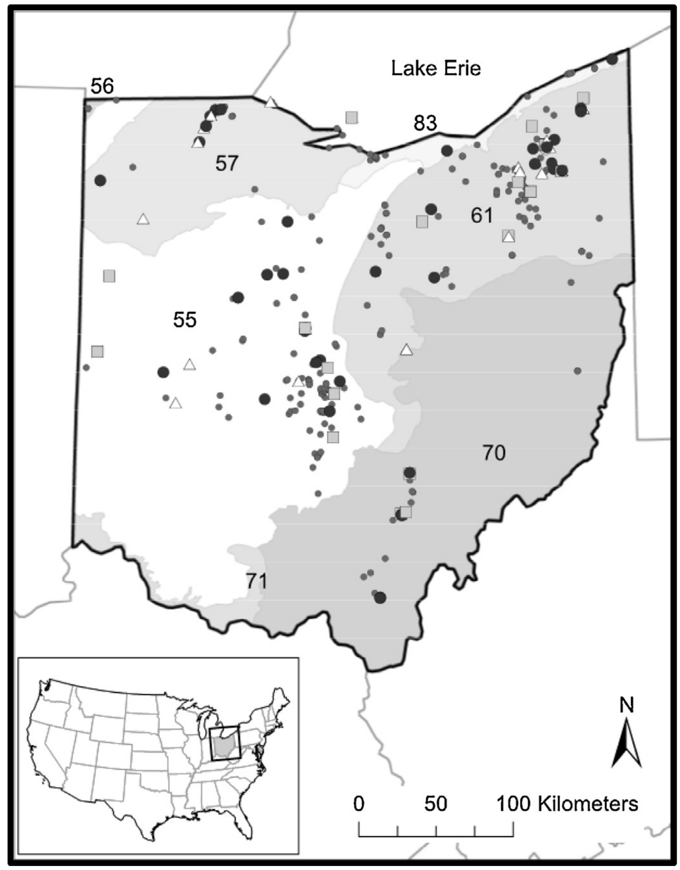

Fig. 1. Location of 252 sampled wetlands in Ohio, USA. The 82 reference-quality wetlands are represented by large symbols and classified by their dominant vegetation type: emergent = white triangles, shrub dominated = grey square, forest = large filled circle. Small dots represent non-reference condition wetlands. Omernik (1987) ecoregions include Eastern Corn Belt Plains (55), Southern Michigan/Northern Indiana Drift Plains (56), Huron/Erie Lake Plains (57), Erie Drift Plain (61), Western Allegheny Plateau (70), Interior Plateau (71), and Eastern Great Lakes and Hudson Lowlands (83).

2. Methods

2.1. General approach

To create modeled vegetation-based IBIs (VIBIs), we mined data that were previously collected from 252 Ohio (USA) wetlands (Fig. 1) between 1999 and 2010 by the Ohio Environmental Protection Agency (OEPA). These wetlands were distributed nonrandomly across Ohio and were previously used to develop classification-based VIBIs for Ohio wetlands (Mack et al., 2000; Mack, 2001, 2004, 2007a,b, 2009). We specifically used these wetlands to ensure maximum comparability with the previously developed classification-based VIBIs. Each wetland had previously been assigned to one of three wetland types: emergent, shrub dominated, or forest. Wetland classification followed Anderson (1982)