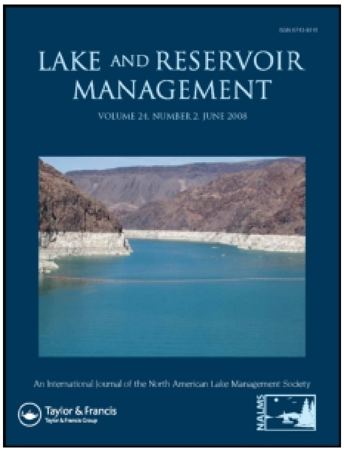

Lake and Reservoir Management
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t905619022

Quantifying regional reference conditions for freshwater ecosystem management: A comparison of approaches and future research needs

Patricia A. Soranno[a]; Tyler Wagner[b]; Sherry L. Martin[ac]; Colleen McLean[cd]; Linda N. Novitski[e]; Cole D. Provence[a]; Allison R. Rober[e]
[a] Department of Fisheries and Wildlife, 13 Natural Resources Building, Michigan State University, East Lansing, MI [b] US Geological Survey, Pennsylvania Cooperative Fish & Wildlife Research Unit, Pennsylvania State University, University Park, PA [c] Department of Geological Sciences, Michigan State University, East Lansing, MI [d] Department of Geological and Environmental Sciences, Youngstown State University, Youngstown, OH [e] Department of Zoology, Michigan State University, East Lansing, MI

First published on: 04 May 2011

To cite this Article Soranno, Patricia A. , Wagner, Tyler , Martin, Sherry L. , McLean, Colleen , Novitski, Linda N. , Provence, Cole D. and Rober, Allison R.(2011) 'Quantifying regional reference conditions for freshwater ecosystem management: A comparison of approaches and future research needs', Lake and Reservoir Management, 27: 2, 138 — 148, First published on: 04 May 2011 (iFirst)

To link to this Article: DOI: 10.1080/07438141.2011.573614

URL: http://dx.doi.org/10.1080/07438141.2011.573614

Lake and Reservoir Management, 27:138–148, 2011

ISSN: 0743-8141 print / 1040-2381 online
DOI: 10.1080/07438141.2011.573614

Quantifying regional reference conditions for freshwater ecosystem management: A comparison of approaches and future research needs

Patricia A. Soranno[1,*], Tyler Wagner[2], Sherry L. Martin[1,], Colleen McLean[3,***], Linda N. Novitski[4], Cole D. Provence[1] and Allison R. Rober[4]**

[1]Department of Fisheries and Wildlife, 13 Natural Resources Building, Michigan State University, East Lansing, MI 48824

[2]US Geological Survey, Pennsylvania Cooperative Fish & Wildlife Research Unit, Pennsylvania State University, 402 Forest Resources Bldg, University Park, PA 16802

[3]Department of Geological Sciences, Michigan State University, East Lansing, MI 48824

[4]Department of Zoology, Michigan State University, East Lansing, MI 48824

Abstract

Soranno PA, Wagner T, Martin SL, McLean C, Novitski LN, Provence CD, Rober AR. 2011. Quantifying regional reference conditions for freshwater ecosystem management: A comparison of approaches and future research needs. Lake Reserv Manage. 27:138–148.

Accurate and robust approaches for quantifying regional reference conditions are critical to the management and restoration of freshwater resources. We considered approaches developed for streams, lakes, or wetlands and for either biological or chemical waterbody features to review 4 common approaches for quantifying regional reference conditions: multimetric, multivariate, landscape-context statistical modeling, and paleolimnology. We focused on the major steps in the decision-making process that led to the most appropriate approach. Based on this synthesis, we argue that there is a need to (1) more explicitly quantify the spatial scale of waterbody variation within and across regions, (2) develop and use predictive classification models in a more explicit fashion to more effectively model this local and regional variation, (3) consider additional metrics with a focus on lakes and wetland responses to both individual and multiple anthropogenic stressors, and (4) continue to develop quantitative approaches to explicitly account for uncertainties in regional reference condition predictions.

Key words: freshwater assessment, paleolimnology, predictive classification models, regional reference conditions, regionalization

Reference conditions are broadly defined as a baseline measure of an ecosystem variable (biological, chemical, or physical attributes) representative of an ecosystem with minimal human influence (Karr 1981, Miller et al. 1988, Stoddard et al. 2006). The concept of a reference condition can be traced back to the development and application of ecological indicators in the early 20th century. The early development of ecological indicators focused on the effects of environmental pollutants on aquatic biota. Since then, ecological indicators and reference conditions have been developed and applied to a wide range of ecosystem types and for a wide range of indicator variables to assess current status and to monitor change over time (Fausch et al. 1984, Karr 1991, O'Connor et al. 2000, Bailey et al. 2004, Niemi and McDonald 2004). To alleviate inconsistencies in the use of the term "reference condition" in the bioassessment literature, Stoddard et al. (2006) proposed that the original concept of reference condition (i.e., with regard to biological integrity) be maintained by changing the term reference condition to "reference condition for biological integrity." This description is useful when specifically discussing the biological condition of ecosystems; however, reference conditions are also developed for a wide range of physical and chemical properties in addition to biological characteristics

*Corresponding author: soranno@msu.edu
**Current address: Department of Geological Sciences, Michigan State University, East Lansing, MI 48824
***Current address: Department of Geological and Environmental Sciences, Youngstown State University, Youngstown, OH 44555

(Goldstein et al. 2002, Bennion et al. 2004, Kershner et al. 2004). Therefore, we use the broad definition of reference condition to also include chemical and physical features of freshwater ecosystems.

The most accurate approach to quantify a reference condition for a freshwater ecosystem is to use data from a time period when the ecosystem was in the "reference" state (i.e., experiencing minimal human impact). However, this is rarely the case, so other methods have been developed that fall into 2 major categories: (1) site-specific reference, or control, sites identified for each system; and (2) regional reference states identified for groups of ecologically similar waterbodies within a region (Hughes et al. 1986). Site-specific reference sites are often used to assess point- and nonpoint-source impacts (Reynoldson et al. 1997), such as when an upstream site is used to represent a nonimpacted reference condition for a downstream portion of a stream. Adequate control sites for each ecosystem that must be assessed are rare; therefore, we focus this paper on approaches to quantify regional reference conditions, defined as the expected condition for a group of ecologically similar waterbodies that are experiencing minimal human disturbance within a predefined region (Hughes et al. 1986, Reynoldson et al. 1997).

A regional reference condition for a given ecosystem type is considered to be representative of similar ecosystems within that region in the absence (or at some minimal level) of human impact (McCormick et al. 2001, Stoddard 2005); however, natural variation among freshwater ecosystems can be large even within small spatial extents (Smith et al. 2003, Wickham et al. 2005, Cheruvelil et al. 2008). Thus, it is desirable to divide the landscape into regions containing ecosystems of similar reference condition. Numerous regionalization schemes delineate the landscape into regions using different features of the landscape. For example, major watershed boundaries (Seaber et al. 1987) are based on drainage patterns, and ecoregions are based on attributes such as geology, natural vegetation, land use and cover, and climate (Bailey 1983, Omernik 1987). This regional approach assumes that the spatial heterogeneity of landscape or climate features constrains important properties of freshwater ecosystems that control variability in freshwater responses. However, several recent studies have shown that much unexplained variation remains within regions, and that additional classification of waterbodies using local hydrogeomorphic (HGM) features is needed to capture the spatial variability within regions (Wang et al. 2003, Snelder and Hughey 2005, Pyne et al. 2007, Cheruvelil et al. 2008, Herlihy and Sifneos 2008, Soranno et al. 2010). Thus, both regionalization and classification (or typology, as stated by Hawkins et al. 2010) of the HGM template of waterbodies must be an integral component of regional reference conditions. We define the HGM template to be any group of the important "natural" hydrologic, morphological, terrestrial, or geological characteristics at any spatial scale that influence waterbody characteristics, such as geology, soils, landscape position, or catchment morphometry.

Quantifying regional reference conditions has become an integral component of freshwater ecosystem assessment and management and is now incorporated into national and international programs, including the US Clean Water Act (CWA) and the EU Water Framework Directive (WFD; Stoddard et al. 2006, Bennion and Battarbee 2007). To effectively implement these mandates, establishment of an expected (i.e., reference) condition to compare the current condition is necessary (Bailey et al. 2004, Stevenson et al. 2004). Many approaches are available for quantifying regional reference conditions (see reviews in Reynoldson et al. 1997, Stevenson and Hauer 2002, Bailey et al. 2004, Bowman and Somers 2005, Stoddard et al. 2006, Hawkins et al. 2010). For example, the application of the reference condition concept in bioassessment has led to the development of comprehensive modeling approaches including the River Invertebrate Prediction and Classification System (RIVPACS; Wright 1995) and related models such as the Australian River Assessment System (AusRivAS; Parsons and Norris 1996), and the Benthic Assessment of Sediment (BEAST; Reynoldson et al. 1995). The vast majority of these applications and tools, as well as the above reviews, have targeted stream macroinvertebrates in minimally disturbed sites. Alternative approaches are needed for freshwater ecosystems in which (1) macroinvertebrates may not be the best measure of reference conditions because they are not good indicators of stressors, and (2) no or few minimally disturbed ecosystems exist.

Our intended audience is researchers and managers working on any freshwater waterbody who have not previously quantified reference conditions. The large number of approaches available to quantify regional reference conditions, each with advantages and disadvantages, can make it difficult to determine which approach is most appropriate under a particular ecological setting and management scenario. A recent review by Hawkins et al. (2010) provides a comprehensive and detailed review of most reference condition approaches developed for streams.

Here, our focus is on all ecosystem types (lakes, wetlands, and streams) as well as the major decision points to determine the appropriateness of different approaches for both biological and chemical responses. Our specific objectives are to (1) summarize 4 methods for estimating regional reference conditions: multimetric, multivariate, landscape-context statistical modeling, and paleolimnology; (2) provide researchers and managers with an outline of steps to quantify regional reference conditions with recommendations as to when certain approaches are most appropriate; and (3) identify research needs for further development and refinement of quantifying reference conditions for ecosystem management.

Approaches to quantify regional reference conditions

The approaches commonly used to quantify regional reference conditions vary depending on data availability, data type, the occurrence of waterbodies representative of reference conditions within a region, and ecosystem type (lake, stream, or wetland). Although other approaches to determine regional reference conditions exist, we highlight the following 4 as representative of the range of possible approaches.

Multimetric

Overview

The multimetric approach integrates multiple characteristics (e.g., species richness or abundance data) of a biological community (e.g., fish, macroinvertebrates, or algae) into a single score that can be used to assess the degree of anthropogenic impact. This approach, often referred to as the "Index of Biological Integrity" (IBI) approach (Karr 1991), has 3 main elements (Stevenson and Hauer 2002): (1) characterization of organisms in reference and disturbed condition sites, (2) calculation of metric scores based on differences in the biological assemblages between reference and disturbed sites, and (3) a multimetric index that combines multiple individual indicators. Reference sites are identified *a priori* using one or more of the following types of information: land use and cover, previous studies or historical data, water quality data, or best professional judgment (Kerans and Karr 1994, Bourdaghs et al. 2006). The multimetric approach has been applied to a variety of taxa in streams (e.g., Kerans and Karr 1994, Rabeni and Wang 2001), lakes (e.g., Lewis et al. 2001, Blocksom et al. 2002), and wetlands (e.g., Wilcox et al. 2002, Uzarski et al. 2005, Wang et al. 2006, Lougheed et al. 2007). The multimetric approach also can incorporate some type of classification of the waterbodies within the region based on local features to minimize natural variability in the indices among the reference and disturbed waterbodies.

Data requirements

The multimetric approach requires species assemblage information and site-level habitat and human disturbance characteristics. Such data should be collected for a number of reference and disturbed sites. Taxonomic identifications at a level adequate to develop a series of structural or functional attributes are used as potential metrics. Individual metrics, such as species richness or relative abundance, are normalized to a standard scale and then summed together to derive a multimetric index (Karr 1991, Kerans and Karr 1994).

Quantifying deviation from reference condition

When comparing reference and test sites, many studies use box-and-whisker plots to examine the amount of overlap between the interquartile ranges of multimetric index values (Barbour et al. 1996, Rabeni and Wang 2001, Blocksom et al. 2002). A test site might be considered impaired if the metric score is less than the 25th percentile for the reference sites within the appropriate region. Various other parametric and nonparametric statistical analyses can also be used to compare the difference between test and reference sites (Kerans and Karr 1994, Rabeni and Wang 2001, Uzarski et al. 2005, Baptista et al. 2007).

Assumptions

This approach assumes the reference sites encompass the full range of natural biological, chemical, and physical variability within the region, and that the samples are unbiased. For the test sites, one assumes the biological community reflects the degree of anthropogenic disturbance exerted on the waterbody and that the biological community integrates the effects of multiple stressors (Stevenson and Hauer 2002). A multimetric index is assumed to provide a more thorough assessment of impairment compared to a single metric.

Multivariate

Overview

The multivariate approach has been widely used for stream invertebrates in the United Kingdom, Canada, and Australia, and for fishes in streams (e.g., Kennard et al. 2006) and lakes (e.g., Tonn et al. 2003). This approach uses species assemblage data similar to the multimetric approach; however, there are fundamental differences in how the data are used (Reynoldson et al. 1997, Bowman and Somers 2005). The 3 most common models for this approach are RIVPACS, AusRivAS, and BEAST (Wright 1995, Parsons and Norris 1996, Reynoldson et al. 1995). All 3 models use multivariate statistical analyses on biological assemblage data to classify reference sites into groups with similar taxonomic composition (Reynoldson et al. 1997). Habitat characteristics of the reference sites are used in a discriminant function analysis to develop a predictive model that defines the expected taxonomic assemblage for a given set of habitat characteristics. These characteristics are then used to assign test sites to one of the reference site classes (Reynoldson et al. 1997).

Data requirements

Multivariate approaches use the same raw data as the multimetric approach. Abundance data are used with BEAST software, while presence/absence data are used for AusRivAS (Reynoldson et al. 1997), and presence/absence data or log abundance data are used with RIVPACS (Wright 2000).

Quantifying deviation from reference condition

Two main methods are used to compare taxa between the test and reference sites (Reynoldson et al. 1997). For RIVPACS and AusRivAS, a test site is compared to reference sites using probability weighting, and a site is determined to be impacted when the number of taxa observed deviates from the number of taxa expected (observed/expected ratio; Reynoldson et al. 1997). For BEAST, a site is determined to be impacted when the abundance and structure of the taxa in the community falls outside of the 90% probability ellipses of the plotted reference site data (Reynoldson et al. 1997).

Assumptions

As with the multimetric approach, the key assumptions are that (1) the biological community reflects the degree of anthropogenic disturbance, (2) the community sample is unbiased, and (3) the sampled reference sites encompass the full range of natural variability in biological, chemical, and physical attributes within the geographic area of interest (Bailey et al. 2004).

Landscape-context statistical modeling

Overview

Landscape-context statistical modeling uses cross-sectional data from many waterbodies along a human disturbance gradient (Seelbach et al. 2002, Baker et al. 2005). This approach has been used in streams and lakes for both biological and chemical response variables. One key difference from the previous 2 approaches is that this approach does not require *a priori* identification of reference sites. Rather, landscape-context statistical models, such as generalized linear models, relate any single response variable to multiple local- and regional-scale HGM and human disturbance variables (Seelbach et al. 2002, Dodds and Oakes 2004, Baker et al. 2005, Soranno et al. 2008). The most common process to calculate site-specific reference conditions is by "hindcasting," in which coefficients for human disturbance variables are set to zero in a regression model that includes all sites and both natural and human disturbance predictors (Baker et al. 2005, Kilgour and Stanfield 2006, Soranno et al. 2008). Thus, the estimated reference conditions account for variability in the HGM features retained in the final model. This approach works best by including a number of waterbodies for which disturbance is already low or zero so that the hindcasting does not occur beyond the range of the data. Nevertheless, when such data are lacking, reference conditions can still be predicted, but the uncertainties in the predicted values will likely be high.

Data requirements

Data requirements are similar to the previous 2 approaches, although the "response" variable is not limited to biota, but can include water chemistry or physical habitat variables. However, these models include more detailed information on HGM and human disturbance characteristics that are commonly available in existing GIS databases. Individual "metrics" are modeled one at a time, thus multiple models can be created for different response variables or metrics of interest.

Quantifying deviation from reference condition

Landscape-context statistical modeling results in an estimated reference condition for each waterbody; thus, comparisons are made between hindcasted reference conditions and current conditions of that particular waterbody (Baker et al. 2005, Soranno et al. 2008). Most applications of this approach provide estimated reference values in some form, but only some studies use the reference value to directly assess impairment. Three main techniques have been used to determine whether an ecosystem has deviated beyond reference state. The first examines whether the human disturbance variables used for model development were statistically significant in the final model. If the variables were not significant, then the ecosystems are assumed to be in reference condition (Baker et al. 2005). The second compares the hindcasted value of a stressor, such as nutrient concentration, with a benchmark created by examining an important biological response to the stressor to determine whether an ecosystem is above or below the benchmark (Soranno et al. 2008). The benchmark is important for determining how far an ecosystem is from a reference state. The third calculates normalized scores for response variables to estimate impaired sites, defined as those with means ± 2 standard deviations from the expected (reference) mean (Baker et al. 2005).

Assumptions

Landscape-context regression models generally assume a linear relationship among waterbody responses and a combination of HGM and human disturbance variables; however, nonlinear models can also be used to incorporate hypothesized or known nonlinear relationships (De'ath and Fabricius 2000, Soranno et al. 2010). In general, the landscape-context regression models depend on several decisions made during the model-building process with assumptions that (1) the model's human disturbances are the dominant disturbances that drive patterns in the waterbody response, (2) the model is valid at low or zero levels of human disturbance, (3) the waterbodies used to build the model constitute a representative and unbiased sample of the entire population of waterbodies within the region, and (4) in regions with no available present day reference sites, the hindcasted conditions from the model are valid. However, the analyst must decide "how far back" the reference state should be modeled. A more appropriate strategy may be to model "best attainable" conditions rather than "minimally disturbed" (Stoddard et al. 2006).

Paleolimnology

Overview

Paleolimnological data from cores have been used to construct regional reference conditions when multiple lakes with core data are classified to allow extrapolation to uncored lakes within the same region (e.g., Cumming et al. 1992, Dixit et al. 1999, Bennion et al. 2004, Battarbee and Bennion 2010, Bennion et al. 2010). Paleolimnological approaches have been used primarily for lentic waterbodies, although some work has been done in running waters (e.g., Gell et al. 2005). Reference conditions using sediment cores can be reconstructed from a range of fossilized indicators representing the flora (e.g., planktonic and benthic diatoms, aquatic plant macrofossils, and pollen) and fauna (e.g., chironomids, ostracods, and cladocerans) in the waterbody, as well as some chemical characteristics of the lake sediments (Cohen 2003). Through the analysis of the top and bottom segments of a sediment core, one can obtain a "before" (reference) and "after" (current) condition of the waterbody. The community composition of fossil diatoms preserved in lake sediments are often used to reconstruct reference levels of total phosphorus concentrations (diatom-inferred total phosphorus [DI–TP]) or lake pH. Next, we describe how the DI–TP can be used to construct regional reference conditions using the following steps: (1) collect predisturbance and present-day DI–TP on as wide a range of lakes as possible (needed for the development of the transfer function models to predict DI–TP if such a model is not available for the region), (2) classify lakes based on natural HGM features (e.g., geology, water depth, alkalinity, or catchment area), (3) quantify DI–TP for predisturbance and present-day time periods, and (4) quantify class-specific reference conditions for all lake classes.

Data requirements

The above approach requires sediment cores from a wide range of waterbodies within a region (i.e., a training dataset), as well as data on HGM features to classify waterbodies. In addition, sediment core depth must be adequate to capture predisturbance conditions along with a wide range of lake types and present-day TP concentrations in the training dataset to develop a robust transfer function.

Quantifying deviation from reference condition

Deviation from reference condition can be estimated for lakes that have been cored and for lakes that have not been cored. For cored lakes, the individual lake's present day DI–TP can be directly compared to its predisturbance DI–TP by comparing whether the change in DI–TP is greater than the root mean square error of prediction (RMSEP) from the transfer function model (Bennion et al. 2004, Leira et al. 2006). Another comparison is the quantification of the squared chord distance (SCD) dissimilarity coefficient of the diatom community composition between predisturbance and present conditions (Bennion and Simpson 2010).

Noncored lakes must first be placed into a lake class of similar lakes based on a classification model of lakes within the study area. A class-level reference condition is quantified for each lake class with at least one sediment core (although more than one core is preferable). Deviation from reference condition is estimated for each noncored lake by comparing its present day TP or DI–TP to the class-level reference conduction using RMSEP to determine if the waterbody has exceeded reference state.

Assumptions

This approach assumes that an adequate range in TP has been included in the training set to develop an accurate transfer function model and that the present flora in unimpacted lakes is similar to the historical unimpacted lakes. In addition, it assumes that a classification can be developed that accurately places lakes into classes that represent their natural reference state prior to disturbance. Where low TP waterbodies are absent due to human disturbance, waterbodies from other similar regions can be used that extend the range of TP for the transfer function models (Bennion et al. 1996). It also assumes that the waterbodies used to develop the transfer functions are similar to the test waterbodies, such that phytoplankton communities in both groups of waterbodies respond similarly to changes in TP. For example, studies have shown that different transfer functions should be developed for shallow and deep lakes (Sayer 2001, Werner and Smol 2005). Finally, this approach can be used only where adequate cores can be obtained from a depositional basin; therefore, this approach cannot be used in many streams, rivers, and some wetlands.

Steps to quantifying regional reference conditions

We synthesized 6 major steps and decision points necessary to quantify regional reference conditions (Fig. 1). Some of these steps have already been well explored for stream macroinvertebrate assemblages and to a lesser degree for

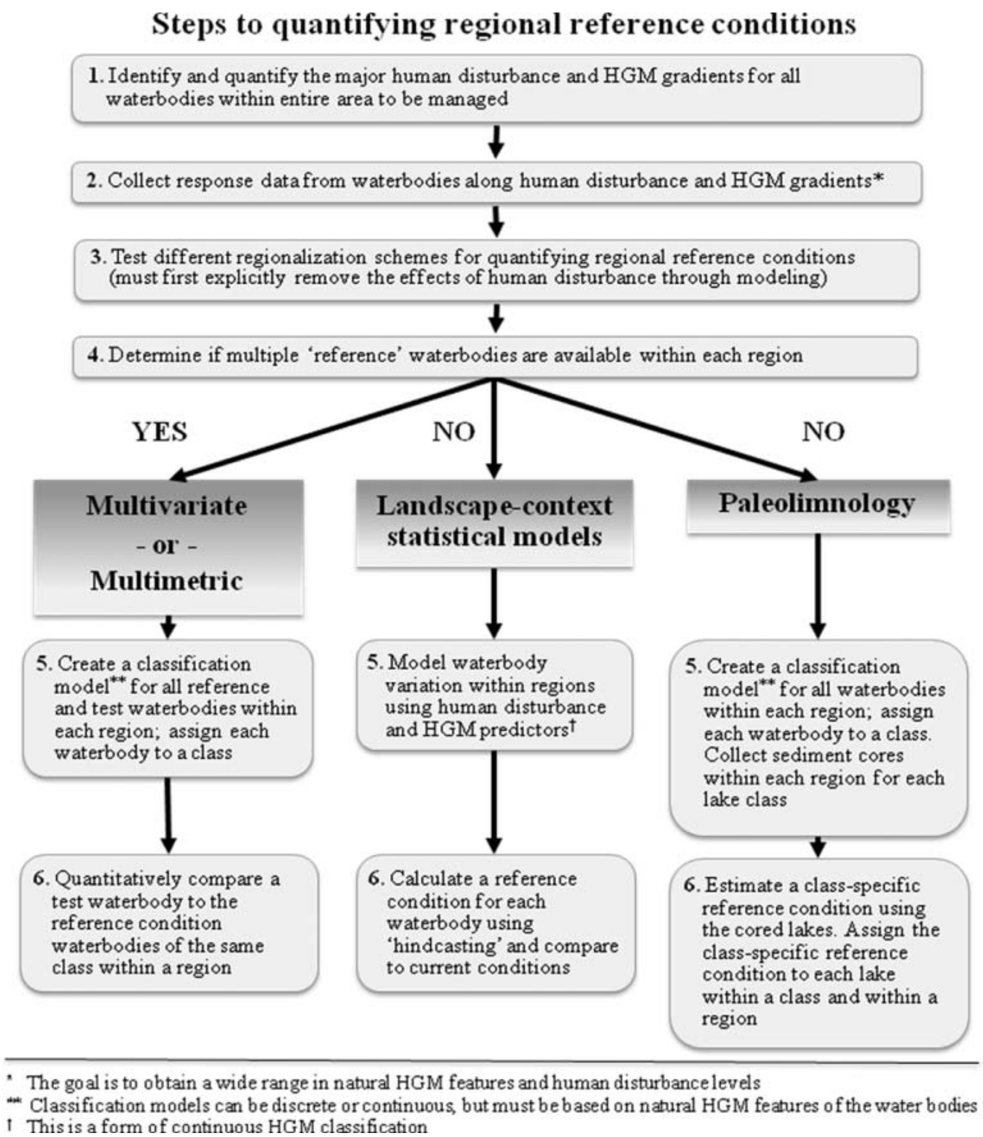

Figure 1.-The steps to quantify regional reference conditions in lakes, streams, or wetlands and the possible approaches that can be used depending on available data. HGM = hydrogeomorphic.

fish assemblages, but not as much for lakes and wetlands or for a range of chemical, physical, and biological characteristics. For all steps, data should be combined into a geographic information system (GIS) database that allows spatial referencing of ecosystems and their surrounding landscape (Johnson and Host 2010).

The first 4 steps are critical to the establishment of regional reference conditions. The first step is to identify and quantify the major human disturbance and natural HGM gradients relevant to the waterbodies under study. These gradients can be identified from previous research on the waterbodies in the region or from the literature from similar regions. The second step is to collect data from a wide range of waterbodies (along both HGM and human-disturbance gradients) for the "response" variable of interest (e.g., fish, macroinvertebrate, or nutrient variables) to quantify regional reference conditions. The temporal component must also be considered; for example, analysts need to decide whether more than one year of data will be collected and across which seasons. The collection of unbiased data is particularly important for this step. When biased data are used, the nature of the bias must be taken into account if possible (Wagner et al. 2008).

The third step is to test and choose the regionalization scheme. An important assumption in the development of regional reference conditions is that waterbodies within a region are more similar than waterbodies across regions.

This assumption must be explicitly quantified and tested to choose the most effective regionalization because recent research shows that regionalizations can differ in accounting for regional variation (Van Sickle and Hughes 2000, Cheruvelil et al. 2008, Yuan et al. 2008). For example, Cheruvelil et al. (2008) showed that lakes within Hydrologic Units (Seaber et al. 1987) were more similar in their water chemistry than lakes within Omernik's Level III ecoregions for 479 lakes in Michigan. Many possible regionalization schemes can be used to quantify regional reference conditions (e.g., different ecoregions or watershed boundaries). Commonly used regionalizations include Omernik's (Omernik 1987) and Bailey's (Bailey 1983) ecoregions, Ecological Drainage Units (Higgins et al. 2005), and watersheds such as Hydrologic Units. To identify which regionalization accounts for the most variation among waterbodies, models should be developed to test among several possible regionalization frameworks. These models must explicitly account for human disturbance effects, which often account for most of the variation explained by ecoregion (Wickham et al. 2005, Cheruvelil et al. 2008), further emphasizing the need to carefully choose the regionalization scheme and decide whether human disturbance variables should be included as part of the regionalization.

The fourth step is to identify waterbodies within each region along a human disturbance gradient to determine whether waterbodies in the reference state (i.e., low or zero human disturbance) are present. This step becomes the major decision point for possible approaches. If an adequate number of candidate reference sites (for each waterbody class) exist within each region, then either the multimetric or multivariate approach can be used. Conversely, if regional reference sites are unavailable, then landscape-context statistical models or paleolimnological methods, where feasible, can be used. In addition, because management agencies must manage potentially thousands of waterbodies within their jurisdictions, site-based approaches may not be a logistically viable option. Within these constraints, few approaches are available to determine regional reference conditions; thus the development of novel or hybrid approaches is important for future research.

The fifth step is to classify waterbodies by the critical HGM variables, which is a key component to all approaches and can be done in a variety of ways, using either discrete or continuous models (e.g., Snelder et al. 2007, Soranno et al. 2010). Because it is not clear whether discrete or continuous models are best for ecosystem management, more research is needed, which we discuss in the following section.

Most efforts to quantify regional reference conditions in the last several decades have focused on the sixth step. For this step in particular there has been much debate regarding the best approach, particularly for multivariate and multimetric approaches in which test sites are quantitatively compared to regional reference sites. The simplified decision tree (outlined in Fig. 1) does not include the rich detail of specific multivariate or multimetric methodologies for steps 1–6, which are discussed elsewhere (e.g., Reynoldson et al. 1997, Bowman and Somers 2005, Stevenson et al. 2009). In contrast, the strategies for the sixth step for landscape-context statistical models and paleolimnological methods have received far less attention. For landscape-context statistical models, a waterbody-specific reference condition is calculated for all lakes using models, whereas for the paleolimnological approach, a class-specific reference condition is estimated using data from sediment cores of representative lakes (Fig. 1).

Finally, we encourage consideration of multiple approaches simultaneously, or hybrid approaches that take advantage of features of more than one approach. For example, hybrid approaches currently being explored incorporate features of more than one of the main approaches reviewed here (Clarke and Murphy 2006, Lavoie et al. 2006, Stevenson 2006). Further, more studies have taken advantage of available GIS databases to improve assessments (Host et al. 2005, Bailey et al. 2007, Collier et al. 2007, Johnson and Host 2010, Soranno et al. 2010). In sum, we recommend a strategy that uses more than one method to compare conclusions using a weight of evidence approach because each method has different strengths and weaknesses.

Future research needs

We highlight 4 research needs to improve the quantification of regional reference conditions. Many of our suggestions relate to the early steps (1–4 in Fig. 1) in the process, prior to the deviations of most approaches; thus, the following suggestions apply to most if not all regional reference condition approaches:

(1) To more explicitly consider the scale of waterbody variation. More research is needed to understand how local and regional variation within waterbodies is partitioned across the landscape and the underlying mechanisms that explain the observed spatial patterns (Wiley et al. 1997, Cheruvelil et al. 2008, Johnson and Host 2010, Soranno et al. 2010). For example, many lowland river sites are highly impacted, so the only unimpacted riverine sites remaining are upland, headwater sites. Thus, we need to understand the landscape context of the systems to assign accurate reference conditions. With a stronger mechanistic understanding of the underlying processes that link local and regional HGM characteristics to waterbody responses, we can develop more robust models for regional reference conditions. Past research has established that both regional and local HGM features influence waterbody variability, and that freshwater

ecosystems are hierarchically organized (Johnson and Host 2010, Soranno et al. 2010). Further, a deeper understanding of this local and regional variation would allow more robust extrapolation and borrowing of data across regions with similar underlying HGM information drivers.

(2) To develop and test Predictive Landscape Classification Models when creating classification models (i.e., Step 5 in Fig. 1). Predictive landscape classification models (Soranno et al. 2010) are based on functional relationships among predictors (in this case, HGM and human disturbance features of waterbodies and their catchments) and response variables such as water chemistry or biology. Such models are advantageous because widely available HGM and human disturbance data in GIS databases, quantified at multiple spatial scales, are available for most if not all waterbodies. In addition, the models are based on principles of landscape limnology that explicitly consider spatial variation of waterbodies at local and regional scales (Soranno et al. 2010). Although predictive landscape classification models are assumed to result in discrete classes, landscape-context statistical models are continuous versions of these types of models and serve similar functions (i.e., they model functional relationships among landscape-context predictors and waterbody responses such as described by Soranno et al. 2008). Thus, the analyst could test whether a continuous or discrete model best classifies natural waterbody variation to gain more accurate estimates of regional reference conditions. The best type of model is not always known *a priori*, so we recommend both be tested. We also recommend further study on multiscaled models (continuous or discrete) in a more comprehensive fashion that considers the hierarchical nature and multiple spatial scales of waterbodies using the best available statistical approaches (Snelder and Hughey 2005, Soranno et al. 2010). Finally, in a recent review, Hawkins et al. (2010) argued that classification models that consider local variables or regionalizations alone should lead to poorer outcomes than site-specific models. The issue should not be to use one or the other approach (classification versus site-specific approaches), however; rather we believe all approaches for quantifying regional reference conditions can be improved through accurate classification models that provide the context for interpreting site-level variation.

(3) To test a wider range of biological and chemical responses to a wider range of stressors. The extensive research on macroinvertebrates in streams has led to the development of reference condition quantification and effective use for management and policy. Although macroinvertebrates are excellent indicators for some stressors in streams, they may not be the ideal indicator for either lakes or wetlands. In addition, macroinvertebrates may not be as well related to other stressors such as hydrologic modification, changes in dissolved organic carbon, or climate change, to name a few. Given that many current threats to freshwater ecosystems occur in combination, we need to examine the responses to multiple stressors rather than individual stressors when considering ideal indicators in waterbodies (Christensen et al. 2006).

(4) To better understand and quantify uncertainties. As with all environmental management problems, quantifying reference conditions and the degree of impairment of test sites must factor in the associated uncertainty (Kelly et al. 2009, Hawkins et al. 2010). Uncertainty in reference condition assessments arise from many sources, including the inherent spatial and temporal variability of freshwater ecosystems, measurement error, prediction uncertainty, and error propagation. Because quantifying regional reference conditions involves several quantitative steps, it is not always clear how errors propagate to the final predictions. Certainly one could try to reduce uncertainties by increasing sample size or using more than one approach to estimate a reference condition. However, it may be more critical to study uncertainty explicitly to identify where the largest sources of uncertainties are and address them directly, as has been done to some degree for selected cases and certain approaches. For example, uncertainty should be quantified in both the definition of the reference condition and the distance of a waterbody from reference condition. Because of the role that reference conditions play in policy making, more research is needed to quantify sources of uncertainty, ways to account for that uncertainty, and the relative influence of uncertainty on the decision making process.

Conclusions

Developing accurate and effective ways to estimate regional reference conditions are critical for managing and rehabilitating freshwater resources. We have synthesized the major steps in the development of regional reference conditions with a focus on the decision points that lead to possible quantitative approaches. We propose that more research is needed to explicitly quantify the spatial scale of waterbody variation across regions; develop and use predictive classification models that incorporate local and regional variation; continue to develop hybrid approaches that take advantage of the best features of existing approaches; develop better approaches for measuring uncertainties in regional reference condition predictions; and finally, to consider additional metrics, particularly in lakes and wetlands, that respond to multiple anthropogenic stressors.

Acknowledgments

We would like to thank the students in the seminar course at Michigan State University in spring semester 2007, “Defining reference conditions in lakes, streams and wetlands:

Approaches and perspectives," for discussions on this topic (T.G. Alwin, D.G. Baas, H.A. Campbell, R.N. Gannon, S.W. Hunt, R.M. Mcninch, M.J. Parsons, M.A. Rinkus and B. Rudolf). Thanks to Katherine E. Webster for many insightful comments and suggestions on earlier drafts that greatly contributed to this manuscript, and 2 anonymous reviewers and the associate editor for helpful suggestions. Finally, P.A.S. and T.W. would like to thank the other members of the Landscape Limnology Research Group (M.T. Bremigan, K.S. Cheruvelil and K.E. Webster) for many discussions on assessing reference conditions in freshwater ecosystems. Use of trade names does not imply endorsement by the federal government.

References

Bailey RG. 1983. Delineations of ecosystem regions. Environ Manage.7:365–373.

Bailey RC, Norris RH, Reynoldson TB. 2004. Bioassessment of freshwater ecosystems: Using the reference condition approach. New York (NY): Kluwer Academic Publishers.

Bailey RC, Reynoldson TB, Yates AG, Bailey J, Linke S. 2007. Integrating stream bioassessment and landscape ecology as a tool for land use planning. Freshwater Biol. 52:908–917.

Baker EA, Wehrly KE, Seelbach PW, Wang L, Wiley MJ, Simon T. 2005. A multimetric assessment of stream condition in the northern lakes and forests ecoregion using spatially explicit statistical modeling and regional normalization. T Am Fish Soc. 134:697–710.

Baptista DF, Buss DF, Egler M, Giovanelli A, Silveira MP, Nessimian JL. 2007. A multimetric index based on benthic macroinvertebrates for evaluation of Atlantic Forest streams at Rio de Janeiro State, Brazil. Hydrobiologia. 575:83–94.

Barbour MT, Gerritsen J, Griffith GE, Frydenborg R, McCarron E, White JS, Bastian ML. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. J N Am Benthol Soc. 15:185–211.

Battarbee RW, Bennion H. 2010. Palaeolimnology and its developing role in assessing the history and extent of human impact on lake ecosystems. J Paleolimnol. DOI 10.1007/s10933-010-9423-7.

Bennion H, Battarbee RW. 2007. The European Union Water Framework Directive: opportunities for palaeolimnology. J Paleolimnol. DOI 10.1007/s10933-007-9108-z.

Bennion H, Battarbee RW, Sayer CD, Simpson GL, Davidson TA. 2010. Defining reference conditions and restoration targets for lake ecosystems using paleolimnology: a synthesis. J Paleolimnol. DOI 10.1007/s10933-010-9419-3.

Bennion H, Fluin J, Simpson GL. 2004. Assessing eutrophication and reference conditions for Scottish freshwater lochs using subfossil diatoms. J Appl Ecol. 41:124–138.

Bennion H, Juggins S, Anderson NJ. 1996. Predicting epilimnetic phosphorus concentrations using an improved diatom-based transfer function and its application to lake eutrophication management. Environ Sci Tech. 30:2004–2007.

Bennion H, Simpson GL. 2010. The use of diatom records to establish reference conditions for UK lakes subject to eutrophication. J Paleolimnol. DOI 10.1007/s10933-010-9422-8.

Blocksom KA, Kurtenbach JP, Klemm DJ, Fulk FA, Cormier SM. 2002. Development and evaluation of the lake macroinvertebrate integrity index (LMII) for New Jersey lakes and reservoirs. Environ Monit Assess. 77:311–333.

Bourdaghs M, Johnston CA, Regak RR. 2006. Properties and performance of the Floristic Quality Index in great lakes coastal wetlands. Wetlands. 26:718–735.

Bowman MF, Somers KM. 2005. Considerations when using the reference condition approach for bioassessment of freshwater ecosystems. Water Qual Res J Canada. 40(3):347–360.

Cheruvelil KS, Soranno PA, Bremigan MT, Wagner T, Martin SL. 2008. Grouping lakes for water quality assessment and monitoring: the roles of regionalization and spatial scale. Environ Manage. 41(3):425–440.

Christensen MR, Graham MD, Vinebrooke RD, Findlay DL, Paterson MJ, Turner MA. 2006. Multiple anthropogenic stressors cause ecological surprises in Boreal lakes. Glob Change Biol. 12(12):2316–2322.

Clarke RT, Murphy JF. 2006. Effects of locally rare taxa on the precision and sensitivity of RIVPACS bioassessment of freshwaters. Freshwater Biol. 51(10):1924–1940.

Cohen AS.2003. Paleolimnology, the history and evolution of lake systems. Oxford (UK): Oxford University Press.

Collier KJ, Haigh A, Kelly J. 2007. Coupling GIS and mulitvariate approaches to reference site selection for wadeable stream monitoring. Environ Monit Assess. 127(1):29–45.

Cumming BF, Smol JP, Kingston JC, Charles DF, Birks HJB, Camburn KE, Dixit SS, Uutala AJ, Selle AR. 1992. How much acidification has occurred in Adirondack region (New York, USA) lakes since pre-industrial times? Can J Fish Aquat Sci. 49:128–141.

De'ath G, Fabricius KE. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. Ecology. 81(11):3178–3192.

Dixit SS, Smol JP, Charles DF, Hughes RM, Paulsen SG, Collins GB. 1999. Assessing water quality changes in the lakes of the northeastern United States using sediment diatoms. Can J Fish Aquat Sci. 56:131–152.

Dodds WK, Oakes RM. 2004. A technique for establishing reference nutrient concentrations across watersheds affected by humans. Limnol Oceanogr-Meth. 2:333–341.

Fausch KD, Karr JR, Yant PR. 1984. Regional application of an index of biotic integrity based on stream fish communities. T Am Fish Soc. 113(1):39–55.

Gell P, Tibby J, Fluin J, Leahy P, Reid M, Adamson K, Bulpin S, MacGregor A, Wallbrink P, Hancock G et al. 2005. Accessing limnological change and variability using fossil diatom assemblages, South-East Australia. River Res Appl. 21:257–269.

Goldstein RM, Wang L, Simon TP, Stewart PM. 2002. Development of a stream habitat index for the Northern Lakes and Forests Ecoregion. N Am J Fish Manage. 22(2):452–464.

Hawkins CP, Olson JR, Hill RA. 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. J N Am Benthol Soc. 29(1):312–343.

Herlihy AT, Sifneos JD. 2008. Developing nutrient criteria and classification schemes for wadeable streams in the conterminous US. J N Am Benthol Soc. 27(4):932–948.

Higgins JV, Bryer MT, Khoury ML, Fitzhugh TW. 2005. A freshwater classification approach for biodiversity conservation planning. Conserv Biol. 19(2):432–445.

Host GE, Schuldt J, Ciborowski JJH, Johnson LB, Hollenhorst T, Richards C. 2005. Use of GIS and remotely sensed data for *a priori* identification of reference areas for Great Lakes coastal ecosystems. Int J Remote Sens. 26:5325–5342.

Hughes RM, Larsen DP, Omernik JM. 1986. Regional reference sites: a method for assessing stream potentials. Environ Manage. 10(5):629–635.

Johnson LB, Host GE. 2010. Recent developments in landscape approaches for the study of aquatic ecosystems. J N Am Benthol Soc. 29(1):41–66.

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

Karr JR. 1991. Biological integrity: a long-neglected aspect of water resource management. Ecol Appl. 1:66–84.

Kelly, M, Bennion H, Burgess A, Ellis J, Juggins S, Guthrie R, Jamieson J, Adriaenssens V, Yallop M. 2009. Uncertainty in ecological status assessments of lakes and rivers using diatoms. Hydrobiologia. 633:5–15.

Kennard MJ, Pusey BJ, Arthington AH, Harch BD, Mackay SJ. 2006. Development and application of a predictive model of freshwater fish assemblage composition to evaluate river health in eastern Australia. Hydrobiologia. 572:33–57.

Kerans BL, Karr JR. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee valley. Ecol Appl. 4:768–785.

Kershner JL, Roper BB, Bouwes N, Henderson R, Archer E. 2004. An analysis of stream habitat conditions in reference and managed watersheds on some federal lands within the Columbia River Basin. N Am J Fish Manage. 24:1363–1375.

Kilgour BW, Stanfield LW. 2006. Hindcasting reference conditions in streams. Am Fish S S. 48:1–18.

Lavoie I, Campeau S, Grenier M, Dillon PJ. 2006. A diatom-based index for the biological assessment of eastern Canadian rivers: an application of correspondence analysis (CA). Can J Fish Aquat Sci. 63(8):1793–1811.

Leira M, Jordan P, Taylor D, Dalton C, Bennion H, Rose N, Irvine K. 2006. Assessing the ecological status of candidate reference lakes in Ireland using paleolimnology. J Appl Ecol. 43:816–827.

Lewis PA, Klemm DJ, Thoeny WT. 2001. Perspectives on use of a multimetric lakebioassessment integrity index using benthic macroinvertebrates. Northeast Nat. 8:233–246.

Lougheed VL, Parker CA, Stevenson RJ. 2007. Using non-linear responses of multiple taxonomic groups to establish criteria indicative of wetland biological condition. Wetlands. 27(1):96–109.

McCormick FH, Hughes RM, Kaufmann PR, Peck DV, Stoddard JL, Herlihy AT. 2001. Development of an index of biological integrity for the Mid-Atlantic Highlands region. T Am Fish Soc. 130:857–877.

Miller DL, Leonard PM, Hughes RM, Karr JR, Moyle PB, Schrader LH, Thompson BA, Daniels RA, Fausch KD, Fitzhugh GA et al. 1988. Regional applications of an index of biotic integrity for use in water resource management. Fisheries. 13:12–20.

Niemi GJ, McDonald ME. 2004. Application of ecological indicators. Ann Rev Ecol Syst. 35(1):89–111.

O'Connor RJ, Walls TE, Hughes RM. 2000.Using multiple taxonomic groups to index the ecological condition of lakes. Environ Monit Assess. 61:207–228.

Omernik JM. 1987. Ecoregions of the conterminous United States. ANN Assoc Am Geogr.77(1):118–125.

Parsons M, Norris RH. 1996. The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. Freshwater Biol. 36(2):419–434.

Pyne MI, Rader RB, Christensen WF. 2007. Predicting local biological characteristics in streams: a comparison of landscape classifications. Freshwater Biol. 52(7):1302–1321.

Rabeni CF, Wang N. 2001. Bioassessment of streams using macroinvertebrates: are the chironomidae necessary? Environ Monit Assess. 71(2):177–185.

Reynoldson TB, Bailey RC, Day K, Norris RH. 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Aust J Ecol. 20:198–219.

Reynoldson TB, Norris RH, Resh VH, Day KE, Rosenberg DM. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. J N Am Benthol Soc. 16(4):833–852.

Sayer CD. 2001. Problems with the application of diatom-total phosphorus transfer functions: examples from a shallow English lake. Freshwater Biol. 46(6):743–757.

Seaber PR, Kapinos FP, Knapp GL. 1987. Hydrologic unit map. Washington (DC): US Government Printing Office; United States Geological Survey Water-Supply Paper 2294.

Seelbach PW, Wiley MJ, Soranno PA, Bremigan MT. 2002. Aquatic conservation planning: Using landscape maps to predict ecological reference conditions for specific waters. In: Gutzwiller K, editor. Concepts and applications of landscape ecology in biological conservation. New York (NY): Springer-Verlag. p. 454–478.

Smith RA, Alexander RB, Schwarz GE. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous United States. Envir Sci Tech. 37(14):2039–3047.

Snelder TH, Dey KL, Leathwick JR. 2007. A procedure for making optimal selection of input variables for multivariate environmental classifications. Conserv Biol. 21(2):365–375.

Snelder TH, Hughey KFD. 2005. The use of an ecological classification to improve water resource planning in New Zealand. Environ Manage. 36(5):741–756.

Soranno PA, Cheruvelil KS, Stevenson RJ, Rollins SL, Holden SW, Heaton S, Torng EK. 2008. A framework for developing ecosystem-specific nutrient criteria: Integrating biological thresholds with predictive modeling. Limnol Oceanogr. 53(2):773–787.

Soranno PA, Cheruvelil KS, Webster KE, Bremigan MT, Wagner T, Stow CA. 2010. Using landscape limnology to classify freshwater ecosystems for multi-ecosystem management and conservation. BioScience. 60(6):440–454.

Stevenson RJ. 2006. Refining diatom indicators for valued ecological attributes and development of water quality criteria. In: Ognjanova-Rumenova N, Manoylov K, editors. Advances in phycological studies. Moscow (Russia): PENSOFT. p. 365–383.

Stevenson RJ, Bailey RC, Harrass MC, Hawkins CP, Alba-Tercedor J, Couch C, Dyer S, Fulk FA, Harrington JM, Hunsaker CT et al. 2004. Interpreting results of ecological assessments. In: Barbour MT, Norton SB, Preston HR, Thornton K, editors. Ecological assessment of aquatic resources: Linking science to decision-making. Society of Environmental Toxicology and Chemistry. p. 85–112.

Stevenson RJ, Hauer FR. 2002. Integrating hydrogeomorphic and index of biotic integrity approaches for environmental assessment of wetlands. J N Am Benthol Soc. 21(3):502–513.

Stevenson RJ, Pan Y, van Dam H. 2009. Assessing environmental conditions in rivers and streams with diatoms. In: Smol JP, Stoermer EF, editors. The diatoms: Applications for the environmental and earth sciences. 2nd ed. Cambridge (UK): Cambridge University Press. p. 57–85.

Stoddard JL. 2005. Use of ecological regions in aquatic assessments of ecological condition. Environ Manage. 34:61–70.

Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecol Appl. 16(4):1267–1276.

Tonn WM, Paszkowski CA, Scrimgeour GJ, Aku PKM, Lange M, Prepas EE, Westcott K. 2003. Effects of forest harvesting and fire on fish assemblages in Boreal Plains lakes: A reference condition approach. T Am Fish Soc. 132:514–523.

Uzarski DG, Burton TM, Cooper MJ, Ingram JW, Timmermans ST. 2005. Fish habitat use within and across wetland classes in coastal wetlands of the five Great Lakes: Development of a fish-based index of biotic integrity. J. Great Lakes Res. 31:171–187.

Van Sickle J, Hughes RM. 2000. Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon. J N Am Benthol Soc. 19(3): 370–384.

Wagner T, Soranno PA, Cheruvelil KS, Renwick WH, Webster KE, Vaux P, Abbitt RJF. 2008. Quantifying sample biases of inland lake sampling programs in relation to lake surface area and land use/cover. Environ Monit Assess 141:131–147.

Wang L, Lyons J, Rasmussen P, Seelbach PW, Simon T, Wiley MJ, Kanehl P, Baker E, Niemela S, Steward PM. 2003. Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, U.S.A. Can J Fish Aquat Sci. 60:491–505.

Wang Y, Stevenson RJ, Sweets PR, Difranco J. 2006. Developing and testing diatom indicators for wetlands in the Casco Bay watershed, Maine, USA. Hydrobiologia. 561:191–206.

Werner P, Smol JP. 2005. Diatom–environmental relationships and nutrient transfer functions from contrasting shallow and deep limestone lakes in Ontario, Canada. Hydrobiologia. 533:145–173.

Wickham JD, Riitters KH, Wade TG, Jones KB. 2005. Evaluating the relative roles of ecological regions and land-cover composition for guiding establishment of nutrient criteria. Landscape Ecol. 20:791–798.

Wilcox DA, Meeker JE, Hudson PL, Armitage BJ, Black MG, Uzarski DG. 2002. Hydrologic variability and the application of index of biotic integrity metrics to wetlands: A Great Lakes evaluation. Wetlands. 22:588–615.

Wiley MJ, Kohler SL, Seelbach PW. 1997. Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. Freshwater Biol. 37:133–148.

Wright JF. 1995. Development and use of a system for predicting the macroinvertabrate fauna in flowing waters. Aust J Ecol. 20(1):181–197.

Wright JF. 2000. An introduction to RIVPACS. In: Wright JF, Sutcliffe DW, Furse MT, editors. Assessing the biological quality of freshwaters, RIVPACS and other techniques. Ambleside (UK): Freshwater Biological Association. p. 1–24.

Yuan LL, Hawkins CP, Van Sickle J. 2008. Effects of regionalization decisions on an O/E index for the US national assessment. J N Am Benthol Soc. 27(4):892–905.