

The reference condition: predicting benchmarks for ecological and water-quality assessments

Authors: Hawkins, Charles P., Olson, John R., and Hill, Ryan A.

Source: Journal of the North American Benthological Society, 29(1) : 312-343

Published By: Society for Freshwater Science

URL: <https://doi.org/10.1899/09-092.1>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

The reference condition: predicting benchmarks for ecological and water-quality assessments

Charles P. Hawkins¹, John R. Olson², AND Ryan A. Hill³

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

Abstract. Benchmarks provide context and are a critical element of all ecological assessments. Over the last 25 y, hundreds of papers have been published on various aspects of ecological assessments, and most of the analyses described in these papers depend on specifying an ecological benchmark for context. Freshwater scientists and managers usually use reference sites (typically sites in natural or least-disturbed condition) to assess the ecological conditions at other sites. Accurate and precise assessments require that assessed sites be matched with appropriate reference conditions. Two general types of approaches have been proposed to predict reference conditions: classifications based on natural environmental settings and models that use continuously variable environmental attributes as inputs. Two types of classifications have been examined: geographic-dependent regionalizations based on general landscape features and geographic-independent typologies that are typically based on combinations of regional and channel features. We examined >1000 papers that addressed some aspect of predicting the reference condition in freshwater ecosystems. We focused on 5 types of benchmarks: ecological, thermal, hydrologic, geomorphic, and chemical. Our review showed that over the last 25 y, researchers have developed increasingly sophisticated methods that can be used to predict reference conditions. Most disciplines have increasingly moved toward site-specific modeling approaches as a way to improve both accuracy and precision of predictions, although typological approaches dominate geomorphic characterizations. Papers published in *J-NABS* have been especially important in advancing and refining methods for predicting ecological benchmarks. Much of the progress made in the science of ecological assessment emerged from research that advanced our understanding of how the spatial and temporal distributions of freshwater biota are related to naturally occurring environmental features and how those relationships can be most accurately and precisely described and predicted. Thus, the performance of ecological assessments is critically linked to how well we characterize freshwater environments, and research in the watershed sciences that addresses predicting thermal, hydrologic, geomorphic, and chemical attributes of freshwater ecosystems has paralleled research focused on predicting biota. We anticipate that knowledge produced from future collaborations between ecologists and watershed scientists coupled with the application of modern modeling techniques will largely determine progress in characterizing and predicting biota–environment relationships and, thus, the accuracy and precision of future ecological assessments.

Key words: reference condition, benchmark, baseline, ecological assessment, landscape classification, ecoregion, regionalization, typology, model, prediction, temperature, hydrologic regimes, geomorphology, water chemistry.

In plain words, Chaos was the law of nature; Order was the dream of man.

Henry Adams (1918)

The objective of an ecological assessment is to measure the status of an ecological resource. Such assessments depend on 2 elements: a measure of the

ecological resource of interest (often expressed in terms of an index) and a benchmark (i.e., a reference condition) from which we can judge if the measured condition of the resource differs from a desired, expected, or previous condition. Without benchmarks, little context exists for interpreting the measured value of an ecological resource because resource states (e.g., number or types of species and their abundances or nutrient concentrations) can vary markedly with natural differences among sites. Understanding the natural variability of ecological

¹ E-mail addresses: chuck.hawkins@usu.edu

² john.olson@usu.edu

³ ryan.hill@usu.edu

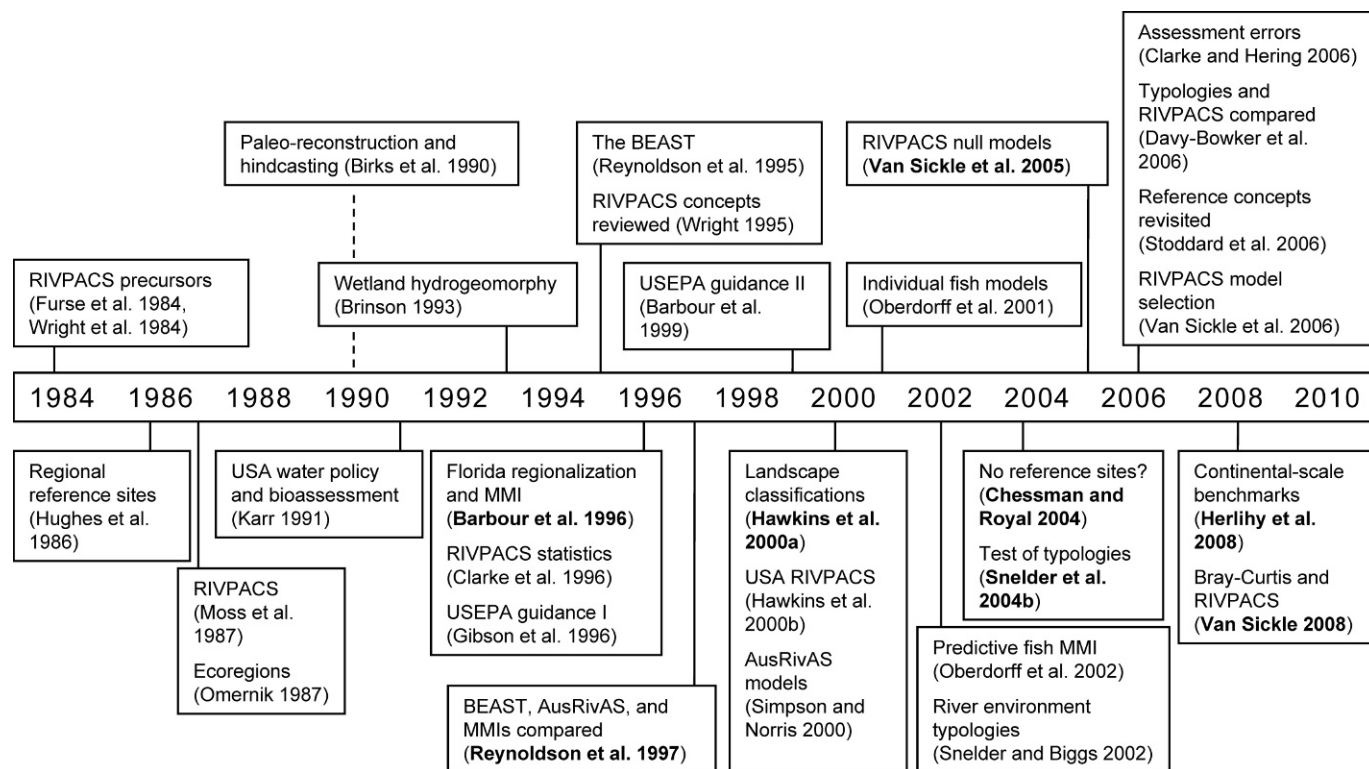


FIG. 1. Important contributions and landmark journal papers and reports describing the development, application, and interpretation of freshwater ecological benchmarks. Dashed lines are used for clarity when a connecting line passes behind a box. Boldface indicates papers published in *J-NABS*. RIVPACS = River InVertebrate Prediction And Classification System, MMI = Multimetric Index, BEAST = Benthic Assessment of Sediment, AusRivAS = Australian River Assessment System, USEPA = US Environmental Protection Agency.

resources and the abiotic conditions associated with that variation is fundamental to the development of sound environmental policy and management directives.

Policy makers and resource managers generally agree that in the context of ecological assessments benchmarks should represent ecological properties associated with naturalness (e.g., Landres et al. 1999, Hering et al. 2003, Stoddard et al. 2006 [Fig. 1]), but no general consensus exists regarding how pristine a condition a benchmark should characterize. In fact, the operational definition of a benchmark is almost always based on some merging of abstract concepts of pristineness, empirical knowledge derived from sites that are the best of what is left and whose quality can differ considerably from historical (pristine), and the ecological conditions desired by human society (which also might differ from pristine conditions). Regardless of the degree of naturalness implied by a specific definition of benchmark, the accuracy and precision of ecological assessments are dependent on the degree to which those benchmarks can be quantified and predicted.

In our paper, we review the development of concepts and methods relevant to predicting reference conditions in stream ecosystems. In doing so, we first summarize the literature published to date that describes different approaches ecologists have used to characterize reference conditions in streams. These approaches generally include methods based on either classification of landscape elements within which streams flow or models that predict site-scale characterizations. We identified 2 types of classifications: geographic-dependent regionalizations that generally describe large, unique, spatially discrete geographic units and geographic-independent typologies that typically describe smaller, spatially repeatable types of landscape settings. We also identified 2 types of models: those models limited to predicting conditions at individual sites (single-site models) and those models that can make predictions of local conditions across a range of environmental settings (site-specific models).

We then review why benchmarks are needed in ecological assessments and discuss a central issue in the application of any assessment: how the inferences

we draw regarding the ecological condition of assessed sites are affected by how we define and estimate benchmark conditions. We next discuss the general role that prediction plays in establishing benchmark conditions, primarily focusing on the 2 major methods ecologists have used to make these predictions: classification and modeling. We conclude by describing how research within the broad discipline of watershed science has helped, and should continue to help, more accurately and precisely predict how environmental conditions vary across natural spatial and temporal gradients and, hence, the most appropriate ecological benchmarks for individual sites.

Methods

To identify relevant literature, we conducted keyword searches in both the Institute for Scientific Information (ISI) Web of KnowledgeSM (1982 to July 2009) and Google Scholar[®]. We also inspected the literature citation sections of the papers compiled from the Web of Knowledge and Google Scholar searches to identify additional relevant papers that these searches missed. We conducted separate literature searches for papers that focused on: 1) establishing and applying ecological reference conditions in all types of freshwater ecosystems and 2) papers in the physical sciences that specifically focused on the prediction of temperature, hydrologic regime, channel geomorphology, or sediment size, and water chemistry in stream ecosystems. We generally restricted our analyses to the primary literature (journals) but included some secondary literature that provided especially important institutional guidance regarding management practices.

For the ecological literature, we first conducted a Web of Knowledge search based on the keywords *reference condition*, *regionalization*, *ecoregion*, *classification*, *typology*, *RIVPACS*, *AusRivAS*, and *BEAST* to identify an initial set of papers that focused on general aspects of the description, prediction, or application of the reference condition in freshwater ecosystems. We then used Google Scholar to identify other candidate publications based on simple combinations of keywords or authors (Google Scholar does not allow Boolean logic searches). We compiled the results of these searches in 3 ways. First, we summarized the number of papers published by individual journals (and their citation rates) that explicitly addressed development or testing of reference condition approaches. Second, we summarized papers that addressed or used some aspect of the reference condition in an ecological assessment by the

type of reference condition approach used: regionalization, typology, single-site, or site-specific models. This latter compilation was substantially larger than the former compilation because it included case studies that did not necessarily develop or assess the performance of the prediction approach used. Last, we summarized the individual papers that appear to have had the most influence on development of reference condition thinking. We identified papers that have been cited ≥ 40 times. The 40-citation cutoff was arbitrary but represented a distinct break from the next most cited paper (18 total citations).

For the physical sciences literature, we conducted 4 separate Web of Knowledge and Google Scholar searches for papers that individually included the keywords *temperature*, *hydrology* or *hydrologic regime*, *channel geomorphology* or *bed sediment*, or *water chemistry* together with the general keywords *stream*, *river*, *lotic*, *reference condition*, *regionalization*, *classification*, *typology*, *model*, and *prediction*. These latter keywords were common to all 4 searches. We identified articles dating back to 1877 that were published in 192 different journals, books, and government documents. We then categorized the information in these papers with respect to whether predictions of reference conditions could be made, and if so, what type of approach to prediction was used.

Results of Bibliographic Analyses of Ecological Papers

Our searches of the ecological literature identified 54 candidate journals (Table 1). Of these journals, 29 had published ≥ 1 paper addressing some aspect of the development or testing of a reference condition approach in freshwater ecosystems (Table 1). Of these journals, *Hydrobiologia*, *J-NABS*, and *Freshwater Biology* (FWB) published 60% of the 184 papers we identified and received 59% of the total citations. Of these 3 core journals, *J-NABS* and *FWB* papers had the highest mean annual citation rates. However, a few papers published in more general ecological journals had higher citations/paper and citations/y (e.g., *Journal of Applied Ecology*, *Australian Journal of Ecology*, *Ecological Applications*, *Global Ecology and Biogeography*, *Ecological Applications*, and *Ecological Modelling*).

Our broad survey of the literature showed that both regionalization and site-specific modeling approaches to establishing reference conditions have been frequently used (Table 2). We associated 413 papers with some type of biological assessment study; 217 studies were based on a classification approach (regionalization or typology), and 187 studies were based on a site-specific modeling approach. Only 9 papers

TABLE 1. Summary of the 184 papers published in 29 journals that addressed some aspects of predicting the ecological reference condition in freshwater ecosystems including number of papers identified in each journal (papers), the total number of citations received by those papers (total cites), the average number of times each paper has been cited (cites/paper), and the average number of citations each paper received/y (cites/year). Twenty-five journals^a did not contain papers that met our criterion for inclusion. These 184 papers have received 4810 citations (Institute for Scientific Information [ISI] Web of Science) as of July 2009.

Journal	Citation statistics			
	Papers	Total cites	Cites/paper	Cites/y
<i>Hydrobiologia</i>	40	436	10.9	2.3
<i>J-NABS</i>	32	978	30.6	4.9
<i>Freshwater Biology</i>	26	1271	48.9	4.9
<i>Environmental Management</i>	12	157	13.1	1.2
<i>Environmental Monitoring and Assessment</i>	9	78	8.7	1.5
<i>Ecological Applications</i>	8	759	94.9	9.7
<i>Ecological Indicators</i>	7	18	2.6	1.3
<i>Archiv für Hydrobiologie/Fundamental and Applied Limnology</i>	7	51	7.3	1.4
<i>Limnologia</i>	5	15	3.0	0.8
<i>Canadian Journal of Fisheries and Aquatic Sciences</i>	4	136	34.0	4.1
<i>Transactions of the American Fisheries Society</i>	4	87	21.8	2.6
<i>Ecography</i>	3	66	22.0	3.9
<i>Water Research</i>	3	49	16.3	2.3
<i>International Review of Hydrobiology</i>	3	34	11.3	2.8
<i>Australian Journal of Ecology</i>	2	294	147.0	10.5
<i>Journal of Applied Ecology</i>	2	73	36.5	11.0
<i>Conservation Biology</i>	2	37	18.5	2.6
<i>Wetlands</i>	2	20	10.0	1.5
<i>Fisheries Management and Ecology</i>	2	20	10.0	5.0
<i>Biological Conservation</i>	2	15	7.5	3.0
<i>Ecological Modelling</i>	1	59	59.0	9.8
<i>Journal of Applied Statistics</i>	1	46	46.0	3.5
<i>Aquatic Conservation: Marine and Freshwater Ecosystems</i>	1	41	41.0	3.7
<i>Water Resources Research</i>	1	22	22.0	3.7
<i>Global Ecology and Biogeography</i>	1	18	18.0	9.0
<i>Environmental Biology of Fishes</i>	1	16	16.0	1.6
<i>Journal of Fish Biology</i>	1	10	10.0	2.5
<i>Journal of the American Water Resources Association</i>	1	4	4.0	0.8
<i>New Zealand Geographer</i>	1	0	0.0	0.0

^a Annual Review of Ecology, Evolution and Systematics; Aquatic Ecology; Aquatic Ecosystem Health; Aquatic Sciences; American Midland Naturalist; Austral Ecology; Ecological Monographs; Ecology; Ecology of Freshwater Fish; Environmental Science and Technology; Fisheries; Fisheries Science; Frontiers in Ecology and the Environment; Journal of Animal Ecology; Journal of Environmental Monitoring; Journal of Great Lakes Research; Limnology; Limnology and Oceanography; Limnology and Oceanography – Methods; North American Journal of Fisheries Management; Oecologia; Oikos; Restoration Ecology; River Research and Applications; Water Environment Research

TABLE 2. Numbers of papers reviewed in each disciplinary area classified by the approach used in setting reference conditions. Individual papers were sometimes assigned to multiple approaches if they compared ways to set reference conditions. The number of different sources (journals, books, reports) for each discipline were: ecology (60), temperature (74), hydrology (53), geomorphology (39), and water chemistry (70).

Approach	Number of papers				
	Ecology	Temperature modeling	Hydrology	Geomorphology	Chemistry
Regionalization	144	0	10	0	19
Typology	73	6	15	49	10
Single site	9	124	0	0	166
Site specific	187	52	122	7	70
Total	413	182	147	56	265

TABLE 3. Eighteen influential papers (i.e., >40 citations) that address reference condition issues in freshwater ecosystems. Papers are ranked by total citations.

Author	Title	Journal	Total citations	Citations/y	Focus
Karr 1991	Biological integrity: a long-neglected aspect of water-resource management	<i>Ecological Applications</i>	451	25.1	General bioassessment
Wright et al. 1984	A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data	<i>Freshwater Biology</i>	274	11.0	Predictive modeling
Reynoldson et al. 1997	The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates	<i>J-NABS</i>	201	16.8	Comparison of indices
Wright 1995	Development and use of a system for predicting the macroinvertebrate fauna in flowing waters	<i>Australian Journal of Ecology</i>	167	11.9	Predictive modeling
Hawkins et al. 2000b	Development and evaluation of predictive models for measuring the biological integrity of streams	<i>Ecological Applications</i>	144	16.0	Predictive modeling
Furse et al. 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	<i>Freshwater Biology</i>	136	5.4	Predictive modeling
Moss et al. 1987	The prediction of the macroinvertebrate fauna of unpolluted running-water sites in Great Britain using environmental data	<i>Freshwater Biology</i>	133	6.0	Predictive modeling
Barbour et al. 1996	A framework for biological criteria for Florida streams using benthic macroinvertebrates	<i>J-NABS</i>	133	10.2	Regionalization
Reynoldson et al. 1995	Biological guidelines for fresh-water sediment based on Benthic Assessment of Sediment (the BEAST) using a multivariate approach for predicting biological state	<i>Australian Journal of Ecology</i>	127	9.1	Predictive modeling
Hughes et al. 1986	Regional reference sites: a method for assessing stream potentials	<i>Environmental Management</i>	117	5.1	Regionalization
Parsons and Norris 1996	The effect of habitat-specific sampling on biological assessment of water quality using a predictive model	<i>Freshwater Biology</i>	105	8.1	Predictive modeling
Hawkins et al. 2000a	Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations	<i>J-NABS</i>	91	10.1	Regionalization
Oberdorff et al. 2001	A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment	<i>Freshwater Biology</i>	78	9.8	Predictive modeling
Dixit et al. 1999	Assessing water quality changes in the lakes of the northeastern United States using sediment diatoms	<i>Canadian Journal of Fisheries and Aquatic Sciences</i>	74	7.4	Historical record
Smith et al. 1999	AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia	<i>Freshwater Biology</i>	74	7.4	Predictive modeling
Clarke et al. 2003	RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers	<i>Ecological Modelling</i>	59	9.8	Predictive modeling
Stoddard et al. 2006	Setting expectations for the ecological condition of streams: the concept of reference condition	<i>Ecological Applications</i>	57	19.0	Concept
Pont et al. 2006	Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages	<i>Journal of Applied Ecology</i>	44	14.7	Predictive modeling

described studies in which the reference condition was based on individual reference sites.

We identified 18 influential papers (Table 3). Among these, Karr (1991; Fig. 1) has been most frequently cited and also has the highest annual

citation rate. Karr (1991) is a general treatment of the need for ecological assessments given environmental protection policy in the USA. This paper does not focus specifically on development of the reference condition, but Karr does discuss the importance of

accurately specifying expected ecological conditions, a topic that received only cursory treatment in Karr (1981), a paper that set the stage for development of multimetric indices (MMIs). Wright et al. (1984; Fig. 1) described the initial analyses in support of the predictive modeling approach to bioassessment (i.e., River InVertebrate Prediction and Classification System [RIVPACS]; Moss et al. 1987; Fig. 1). Wright et al. (1984) is the single most frequently cited paper that specifically addresses prediction of the local biota expected under different natural environmental conditions. Of the remaining 16 papers, 10 built on Wright et al. (1984) and described aspects or refinements of the modeling approach, 3 explored aspects of the regionalization or typology approaches to setting reference expectations (Hughes et al. 1986, **Barbour et al. 1996**, **Hawkins et al. 2000a**¹; Fig. 1), 1 treated the overall challenges in identifying reference conditions and applying them in a uniform manner (Stoddard et al. 2006), 1 paper described how paleolimnological data could be used in lakes to identify historical conditions for individual sites (Dixit et al. 1999), and 1 compared the performance of different types of biotic indices (**Reynoldson et al. 1997**; Fig. 1). Reynoldson et al. (1997) is the single most cited paper on bioassessment topics published in *J-NABS*. Its main focus was on comparing indices, but the paper was critical in establishing the need to tease apart the potentially confounding effects on bioassessments of how the reference condition is predicted from the ecological index used.

J-NABS papers have been an eclectic contribution to the development of reference condition ideas (Fig. 1, Table 3). Three of the 18 influential papers identified in Table 3 were published in *J-NABS*, but we identified 5 other *J-NABS* papers that either have influenced reference-condition ideas or that we think probably will do so in the future. Barbour et al. (1996) built on the pioneering paper of Hughes et al. (1986) in describing the implementation of a regionalization approach for setting biological expectations. Eleven years after Hughes et al. (1986), Reynoldson et al. (1997) described the first rigorous evaluation of the performance of 2 different index types (both of which depend on comparison with reference conditions). Three years later, Hawkins et al. (2000a) synthesized the results of several papers published in a special issue of *J-NABS* on landscape classification (volume 19, issue 3) that evaluated how well regionalizations and typologies accounted for natural biological variation among streams. Snelder et al. (2004b; Fig. 1) followed with an evaluation of perhaps the

most comprehensive ecologically based typology of streams developed to date. In the same year, Chessman and Royal (2004; Fig. 1) described an approach to estimating reference conditions when no reference sites were available. Van Sickle (**Van Sickle et al. 2005**, **Van Sickle 2008**; Fig. 1) provided new tools to enhance the development, interpretation, and application of RIVPACS-type models. Most recently, Herlihy et al. (2008; Fig. 1) built on the synthetic work of Stoddard et al. (2006) and others to describe the challenges implementing a consistent approach to setting reference expectations at a continental scale. As evident in Fig. 1, there has been considerable activity over the last 25 y in developing and refining approaches to describe and predict ecological benchmarks in freshwater ecosystems. Figure 1 provides context for our more detailed review below of the contributions that these and other researchers have made to the theory and practice of ecological assessments.

Developing and Applying Benchmarks in Ecological Assessments

The need for control and replication in ecological assessments

Why are benchmarks needed in ecological assessments? Simply put, a benchmark should serve the same purpose as a control treatment in a manipulative experiment. Thus, benchmarks are one element of a study design that helps resource managers draw scientifically credible inferences regarding the ecological condition of the resources they manage. Like the controls used in a good manipulative experiment, benchmarks should accomplish 2 main tasks: 1) control for the effects of factors other than the ones being studied (e.g., natural factors vs anthropogenic stressors) and 2) provide sufficient replication to estimate the range of variation in the values of an ecological index that is associated with both natural variability and sampling variability. The extent to which we accomplish these 2 tasks will greatly affect the robustness of inferences regarding the ecological condition of a water body.

Experimental control and replication are fundamental requirements in any ecological assessment, but the specific way these design elements have been applied has been strongly influenced by practical constraints imposed by a heterogeneous landscape and by the statistical sophistication of practitioners. Tightly matched controls can be applied when conducting manipulative field experiments designed to evaluate the potential effects of one of more types of stressor or levels of a stressor on an ecological variable of interest (see review by Cooper and

¹ Boldface indicates paper was published in *J-NABS*

Barmuta 1993), but such highly controlled approaches are difficult if not impossible to use for post hoc assessments of the effects of environmental alterations that have already occurred (potentially over a period of ≥ 100 y) at individual sites (e.g., Underwood 1994). The need to conduct post hoc assessments at individual sites prompted development of approaches to experimental control that were designed to characterize better the sources and magnitude of natural variability that occur at the scale of entire ecosystems (e.g., Landres et al. 1999).

Two general approaches to matching control sites with an assessed site are frequently used in freshwater ecological assessments. The 1st approach requires the identification of nonperturbed reference sites that are assumed to be sufficiently similar to one another and the predisturbance condition of the assessed site that they can be used as replicates of fixed controls. The 2nd approach assumes that natural environmental and ecological variability among reference sites will be so great that their direct use as fixed replicates could mask human-caused effects (see following paragraph). Instead of using reference sites as fixed replicates, this approach uses biota–environment relationships derived from the reference sites to predict the most likely ecological reference condition at any individual assessed site (analogous to regression).

Replication is a critical component in any type of assessment, but its use in ecological assessment and monitoring has differed in important ways from its use in classic controlled experiments (i.e., analysis of variance [ANOVA]-type designs). Replication in controlled experiments assumes treatments are fixed, and replicates are used to estimate sampling error. That estimate of sampling error is then used to test statistically the likelihood that the observed control and treatment means are different. In ecological assessments of individual sites, the null hypothesis that control and treatment means are identical is not appropriate because there is only one treatment observation. It might be tempting to use replicate samples taken within an assessed site to conduct such a test, but the variance associated with within-site replicates typically will be smaller than the variance observed among reference sites (within-site + among-site variation), thereby invalidating the assumption of equal variance and, hence, the statistical test. If used in such tests, the replicate samples taken within a site represent pseudoreplicates (*sensu* Hurlbert 1984), and their use in a *t*-test (or similar test) would inflate the likelihood of observing a statistically significant difference in means. The appropriate hypothesis when assessing individual sites is whether the condition observed at that site is outside the range

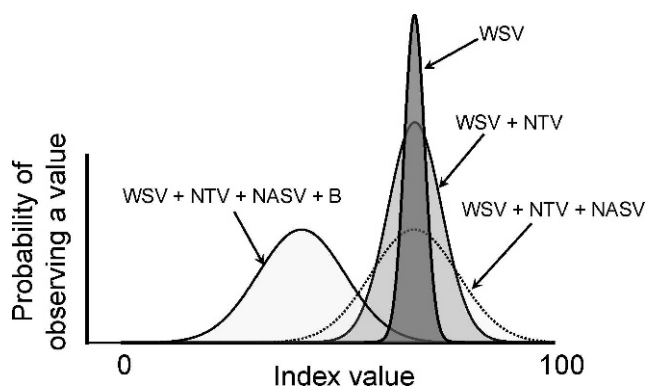


FIG. 2. Components of variability in a hypothetical ecological index associated with estimating the reference condition. WSV = within-site sample variation, NTV = natural temporal variation, NASV = natural among-site variation, B = bias associated with systematic measurement or prediction error.

of index values expected among an appropriate set of reference sites (see McBride et al. 1993, Kilgour et al. 1998, Smith et al. 2005, Bowman and Somers 2006 for appropriate statistical approaches).

Precision, bias, and inferences in ecological assessments

Reliable inferences of ecological condition assume that predictions of the reference condition are both acceptably precise and unbiased. Ideally, the only variation associated with estimating the reference condition is that associated with random sampling variability, i.e., the variation in an ecological attribute that would occur among replicate samples taken at an individual site at the same time. In reality, sampling variability is usually confounded with 2 additional types of variation in ecological assessments: 1) natural, systematic variation that occurs over time at individual sites and that exists among the populations of reference sites used as spatial replicates and 2) systematic variation associated with prediction error. Because all natural ecosystems are inherently dynamic, they exhibit a range of conditions associated with temporal variation in both abiotic and biotic forces (Fig. 2). Short-term temporal variation can be associated with either predictable seasonal processes, such as life-history schedules of component species or effects of seasonal variation in climate, or stochastic events, such as individual floods or periods of extreme drought. Because of this seasonal variation in assemblage structure, and thus, metric or index values, assessments based on comparisons of samples taken in one season with reference expectations derived from other seasons will lead to biased predictions and incorrect inferences (e.g., Linke et al.

1999, Reece et al. 2001, Feio et al. 2006). Longer-term year-to-year differences in environmental conditions associated with climate can directly or indirectly affect annual variation in species composition, richness, or relative abundances via effects on the recruitment success and survivorship of individual species (McElravy et al. 1989, Bradley and Ormerod 2001, Scarsbrook 2002, Bêche and Resh 2007).

The effects of natural variation typically have been treated in different ways. Controlling for seasonal variability can be done, in part, by specifying a relatively short period within which samples are collected (e.g., an index period; Barbour et al. 1999; Fig. 1) or by pooling multiple samples taken across seasons to estimate the annual ecological conditions that characterize a site (e.g., Wright 2000). Year-to-year variation typically is considered to be an important component of the range of natural variability that characterizes a site. This type of variation is fundamentally different from random sampling variability in that it is clearly a consequence of reasonably well-understood causal processes. It is also an important component of ecological variation that should be used to assess whether sites are in reference condition (Landres et al. 1999). Spatial variation among reference sites is an outcome of the pervasive and multidimensional environmental heterogeneity that exists naturally across all landscapes and waterways. Thus, such sites can be, at best, coarse replicates of one another. In regionalization or typology approaches (i.e., any type of landscape classification), this variation is explicitly or implicitly assumed to represent part of the range of natural variation associated with a given water body type, and some lower value of that distribution of values is used as a benchmark below which an assessed site would be considered to be in nonreference condition. In site-specific approaches, this type of variation represents a potentially important confounding factor that can reduce estimates of precision and, therefore, should be taken into account by modeling.

Systematic prediction bias occurs as a consequence of nonrandom prediction errors and can occur in any type of prediction scheme. For example, given that we know that the distribution of biota varies strongly with stream size, a classification of sites that treats small and large reference streams as equivalent would probably either over- or underpredict the true reference value for any individual site. Such biased predictions also can occur in models that are developed to control for naturally occurring environmental variation among sites. Predictions at some sites will be inaccurate if the models imperfectly describe environment–biota relationships.

These issues of prediction errors have attracted increasing scrutiny as their existence and magnitude have become more apparent over the last ~10 y. Most analyses of assessment error have focused on aspects of precision (Clarke 2000, Hawkins et al. 2000b [Fig. 1], Ostermiller and Hawkins 2004, Clarke and Hering 2006 [Fig. 1], Davy-Bowker et al. 2006 [Fig. 1], Nichols et al. 2006, Van Sickle et al. 2007, Stribling et al. 2008, Mazor et al. 2009). Recent research has been directed toward documenting and understanding the magnitude of systematic prediction errors (e.g., Ostermiller and Hawkins 2004, Hawkins 2006, Cao et al. 2007, Hargett et al. 2007, Rehn et al. 2007, Ode et al. 2008). These studies have made it increasingly clear that substantial variation in reference-site biota and the indices derived from them can be associated with natural gradients, even after modeling or adjusting for the effects of natural environmental variation. For example, as much as 30% of the variation in reference-site index values was associated with natural environmental gradients in the USA's recent National Wadeable Stream Assessment (USEPA 2006) (Table 4). The magnitude of this variation implies that assessment errors associated with biased predictions might be frequent and that the issue of bias needs much more attention than it has received.

A brief history of approaches used to estimate ecological benchmarks

Ecological benchmarks can be estimated in at least 4 ways: 1) use of the historical record, 2) extrapolation or interpolation from extant reference-quality sites of appropriate type, 3) hindcasting based on models that describe current stressor–indicator relationships, and 4) prediction derived from mechanistic models that describe how natural processes influence an ecological attribute of interest. Matching of existing reference-quality sites with assessed sites provides a means of estimating aspects of both assemblage composition (what taxa should occur) and structure (what abundances should occur) and is, by far, the main approach used in modern ecological assessments. Matching can be done in several ways, of which landscape regionalizations, typologies, and site-specific predictions are most common. Use of the historical record, hindcasting, and mechanistic modeling are much less well developed (discussed below).

Landscape regionalizations.—In an effort to produce ecological assessments broadly based on the design elements of control and replication, Hughes et al. (1986) advocated use of a regional reference site

TABLE 4. Results of Random Forests regression modeling assessing associations between 2 types of biological index (multimetric index [MMI] and observed/expected taxa ratio [O/E]) and natural variables. Index values were regressed against 7 natural variables (DOY = day of year sample was taken, TMEAN = mean long-term annual air temperature, PRECIP = mean annual long-term precipitation, LAST-0 = mean long-term average day of the last freeze, WSAREA = watershed area, CARB = % of watershed with carbonate geology, and TRANGE = difference between mean long-term average maximum and minimum air temperatures). MMI values were standardized by ecoregional means to have within-region means of 1. Variance = % of variance in index values associated with the natural gradients. Wadeable Streams Assessment (WSA) ecoregions (USEPA 2006, Herlihy et al. 2008) are Coastal Plains (CPL), Northern Appalachian Mountains (NAP), Northern Plains (NPL), Southern Appalachian Mountains (SAP), Southern Plains (SPL), Temperate Plains (TPL), Upper Midwest (UMW), Western Mountains (MTN), and Western Xeric (XER).

WSA ecoregion	Index type	N	Variance	Ranked variable importance					
CPL	O/E	15	0.0						
	MMI	15	0.0						
NAP	O/E	54	0.0						
	MMI	54	2.3	TMEAN	WSAREA	CARB	PRECIP	LAST-0	
NPL	O/E	17	5.4	PRECIP	DOY	TMEAN			
	MMI	17	0.0						
SAP	O/E	110	0.0						
	MMI	110	30.1	WSAREA	LAST-0	TMEAN	DOY	TRANGE	
SPL	O/E	20	0.0						
	MMI	21	17.3	LAST-0	PRECIP	TMEAN	TRANGE	CARB	
TPL	O/E	37	6.7	WSAREA	LAST-0	TMEAN	DOY	CARB	
	MMI	37	22.4	LAST-0	DOY	PRECIP	TMEAN	WSAREA	
UMW	O/E	12	0.0						
	MMI	12	0.0						
MTN	O/E	119	4.6	DOY	TRANGE	LAST-0	TMEAN	PRECIP	
	MMI	120	7.6	PRECIP	TMEAN	DOY	LAST-0	TRANGE	
XER	O/E	32	0.0						
	MMI	32	13.4	DOY	TMEAN	LAST-0	WSAREA	TRANGE	
ALL	O/E	416	4.1	TMEAN	DOY	LAST-0	PRECIP	WSAREA	
	MMI	418	17.3	PRECIP	TMEAN	LAST-0	WSAREA	DOY	

approach. In the regional reference site approach, some set of ecologically least-disturbed sites is identified within a relatively homogeneous terrestrial landscape (the controls), and the range of biotic attributes observed across this population of sites (the replicates) is used to estimate the range of natural variability that presumably would occur at an individual test site (a site of unknown condition) assuming it was in reference condition. If the value of an ecological index falls outside a threshold value determined from the distribution of reference site values, the site would be inferred to be in nonreference condition. Inferences regarding the condition of a test site initially were based on somewhat arbitrary and simple thresholds (e.g., the 10th or 25th percentile of reference-site values), but formal statistical tests (e.g., Kilgour et al. 1998, **Bowman and Somers 2006**) can be applied to such data to determine more rigorously whether an assessed site is in reference condition.

Hughes et al. (1986) proposed the use of ecoregions (Omernik 1987; Fig. 1) as landscape strata within which assessed sites would be compared with reference sites. The general rationale for using ecoregions as

a basis for predicting aquatic biota was based, in part, on Hynes' (1975) view that streams are a product of the terrestrial ecosystems they traverse (see fig. 1 in Hughes et al. 1986). Hughes et al. (1986) suggested that if ecoregions are defined in terms of their climate, topography, geology, soil, and vegetation (Omernik 1987), then their streams should differ in water chemistry, flow regime, habitat structure, and food sources. In turn, these proximate environmental features should strongly influence the biotic character of streams. Thus, ecoregions were viewed as a way to partition (control for) the collective effects of the most important natural factors that control the distribution and abundance of aquatic biota. An important property of the ecoregion approach to regionalization is that the spatial units used for prediction are generally all unique, i.e., ecoregions are a type of geographic-dependent classification in which each class typically has only one member. Thus, no true replicates of ecoregions exist, and assessing the transferability of knowledge gleaned from one ecoregion (or set of ecoregions) to others is problematic.

The regional reference site approach developed nearly in parallel with a major effort by the US

Environmental Protection Agency (EPA) to define and describe ecoregions within the USA (Omernik 1987), and subsequently, an ecoregional reference site approach was widely promoted by the EPA as a method for controlling, in whole or part, natural ecological variation when conducting biological assessments in the USA (Gibson et al. 1996 [Fig. 1], Barbour et al. 1999). A regionalization approach to defining reference conditions was subsequently applied elsewhere (e.g., European Commission 2000, Moog et al. 2004, Verdonschot and Nijboer 2004). Studies published in *J-NABS* did not lay the original foundations for the use of such landscape regionalizations in ecological assessments, but several papers that documented refinement of the approach and case studies of its application appeared in *J-NABS* (e.g., Lenat 1988, 1993, Barbour et al. 1996, Pan et al. 1996, 2000, Gerritsen et al. 2000, Maxted et al. 2000, Van Sickle and Hughes 2000, Waite et al. 2000).

Typologies.—The eventual recognition that almost any type of regionalization (ecoregion, drainage basin, political boundaries) was associated with about the same, and relatively low, amounts of biotic variation among sites (Hawkins et al. 2000a, Marchant et al. 2000, Sandin and Johnson 2000, Van Sickle and Hughes 2000, Waite et al. 2000, Herlihy et al. 2006, 2008, Herlihy and Sifneos 2008) implied that reassessment was needed of assumptions regarding which specific proximate features most strongly influence the distribution of aquatic biota (e.g., Power et al. 1988, Hawkins et al. 1997, Poff 1997, Mykrä et al. 2007) and what landscape and catchment attributes best predicted those specific features. This recognition anticipated development and testing of more finely resolved, geographic-independent classifications for streams based on spatially repeatable ecotypes or landscape typologies (e.g., Brierley and Fryirs 2000, Waite et al. 2000, Snelder and Biggs 2002 [Fig. 1], Fryirs 2003, Balestrini et al. 2004, Munné and Prat 2004, Snelder et al. 2004a, b, Higgins et al. 2005, Chessman et al. 2006, Seelbach et al. 2006, Schmitt et al. 2007, Orr et al. 2008). Although not initially a main thrust of stream classification research, such typological thinking had an earlier start in the hydrogeomorphic classification of wetlands (e.g., Brinson 1993 [Fig. 1], Brinson and Rheinhardt 1996, Shaffer et al. 1999). The mandate of the European Union Water Framework Directive (WFD) that ecological assessments must be comparable across member nations prompted several studies that assessed the effectiveness of different typologies in bioassessment (e.g., Lorenz et al. 2004, Verdonschot and Nijboer 2004, Dodkins et al. 2005, Ferréol et al. 2005, Aroviita et al. 2008, Turak and Koop 2008), especially regarding how

to develop optimal typologies based on only those catchment and waterbody features that most strongly influence biota (e.g., Snelder et al. 2004b, 2007, Dodkins et al. 2005, Snelder and Hughey 2005, Heino and Mykrä 2006, Sánchez-Montoya et al. 2007, Aroviita et al. 2008). However, tests of the effectiveness of the typological approach in partitioning natural ecological variation have not produced encouraging results (Davy-Bowker et al. 2006, Inglis et al. 2008).

Papers published in *J-NABS* were among the first to explore the utility of a priori classifications other than ecoregions for bioassessment purposes (e.g., review by Hawkins et al. 2000a, Waite et al. 2000), and general ecological syntheses describing the ecological knowledge that underlies the conceptual foundations on which typologies are based were published in *J-NABS* (e.g., Power et al. 1988, Poff 1997). However, most of the recent advances in development and testing of typologies (see above) have been published elsewhere, although the pioneering analyses of Snelder et al. (2004b) is an important exception.

Site-specific prediction of reference conditions.—During the same time period that the regional reference site approach was being developed, an alternative approach to establishing reference conditions was being developed and refined in Great Britain. This approach was designed to derive nearly site-specific reference expectations by adjusting index values for the effects of natural environmental attributes that can vary both among and within regions. Moss et al. (1987) (see also Furse et al. 1984 [Fig. 1], Wright et al. 1984, Wright 1995 [Fig. 1], Clarke et al. 1996 [Fig. 1], 2003, Moss et al. 2001) developed statistical procedures (RIVPACS) to predict taxon-specific probabilities of detection (*P*) at different sites from naturally occurring environmental features. The statistical models on which RIVPACS-type systems are based are calibrated with observations made at many reference-quality sites, and the models relate taxon occurrences to multiple environmental gradients. In effect, this approach simultaneously models the niche relationships of many taxa.

In the original RIVPACS and its many refinements and derivatives (e.g., Reynoldson et al. 1995 [Fig. 1], Parsons and Norris 1996, Smith et al. 1999, Hawkins et al. 2000b, Simpson and Norris 2000 [Fig. 1], Joy and Death 2002, Davy-Bowker et al. 2006, Hawkins 2006, Kokeš et al. 2006, Feio et al. 2009, Poquet et al. 2009), no a priori classification of reference sites is used (unlike the regional reference or typology approaches). Instead, RIVPACS-type models predict *P* by first classifying reference-quality streams based on their taxonomic composition or assemblage structure (not

their landscape or channel environmental attributes) and then applying a multivariate predictive model (e.g., linear discriminant functions models [LDMs]) to predict class membership from environmental variables that are generally invariant (unresponsive to anthropogenic activity). Because LDMs predict the probabilities that new observations belong to each of the different classes, the P_s at a site can be estimated by weighting the frequencies of occurrence of each taxon within each group by these probabilities of group membership. This weighting allows prediction of P continuously across the entire predictor variable space encompassed by the reference sites. In essence, the models interpolate the likelihood of collecting different taxa at a site from the reference-site groupings. Excessive extrapolation is prevented by determining if the predictor space occupied by an assessed site is outside that described by the reference sites (e.g., Moss et al. 1987, Clarke et al. 2003).

In the RIVPACS approach, taxon-specific P_s are used to derive ecological quality ratios (EQRs) (Clarke et al. 1996, 2003). For example, O/E_{taxa} is calculated by first summing the P_s to estimate the number of taxa expected (E) in a sample. The O/E_{taxa} index is the proportion of those specific taxa predicted to occur in the sample (i.e., E) that are observed (O). When used in conjunction with other ecological information, such as tolerance values, the P_s also can be used to estimate EQRs that assess if an observed tolerance-based biotic index value is different from expected (e.g., the Biological Monitoring Working Party [BMWP] indices; Clarke et al. 2003). In both cases, inferences regarding the ecological condition of a site are based on estimates of modeling error, i.e., the distribution of model residual values. In theory, the distribution of O/E values calculated at reference sites should have a mean of 1, and the variance of this distribution is a direct measure of the precision of a model. If these values are normally distributed, they allow the same type of statistical tests of whether an assessed site is in reference condition or not as previously mentioned (e.g., Kilgour et al. 1998, **Bowman and Somers 2006**). As discussed above, matching the spatial and temporal extents over which indices are calculated and how reference-site variability (i.e., modeling error + natural variation) is estimated should reduce Type I and II statistical errors of inference. Although EQRs typically have been estimated following the traditional RIVPACS modeling approach, EQRs (including O/E_{taxa}) also can be calculated from an a priori classification of sites (e.g., Aroviita et al. 2008, 2009). Moreover, in the null model approach developed by Van Sickle et al. (2005) as an aid in evaluating the performance of RIVPACS-type models, O/E is calcu-

lated by assuming that all reference sites belong to a single regional class.

As with the regional reference site approach, foundational research in developing site-specific reference conditions was not published first in *J-NABS*. However, *J-NABS* has published the 2nd-most papers that the ISI Web of Knowledge (as of July 2009) identified from the keywords RIVPACS or AusRivAS (Australian River Assessment System) (15 of 106 papers). Moreover, a *J-NABS* paper (**Reynoldson et al. 1997**) has been cited more often (201 citations) than any of these other papers. Many of the *J-NABS* papers on RIVPACS-type topics focus on ways to improve predictions (**Van Sickle et al. 2005, 2006** [Fig. 1], **Cao et al. 2007, Mykrä et al. 2008b**), better estimate and interpret prediction errors (**Ostermiller and Hawkins 2004, Van Sickle et al. 2005, Yuan 2006, Ode et al. 2008, Yuan et al. 2008**), or derive more sensitive indices based on model outputs (**Van Sickle 2008**). Other papers describe case studies that illustrate the utility and potential of a site-specific reference condition approach (**Sloane and Norris 2003, Hose et al. 2004, Cao et al. 2007**).

In contrast, *J-NABS* has not led in the publication of RIVPACS-type assessments in which methods other than biotic-based site classification coupled with LDMs are used to predict probabilities of capture. Some recent work has focused on predictive schemes that do not require a preclassification step. This work appears to have been stimulated primarily by the argument that the classification step used in RIVPACS might not minimize prediction errors. For example, Joy and Death (2003) and Olden (2003) explored the effectiveness of linking the outcomes of many taxon-specific models to estimate P_s . Others have used artificial neural networks (ANNs) to jointly predict multitaxa P_s directly from environmental data (e.g., Joy and Death 2005, Olden et al. 2006). Still others have developed approaches that identify that subset of reference sites that are best physically matched with a test site of interest—a type of site-specific typology (e.g., Linke et al. 2005, Prins and Smith 2007, Chessman et al. 2008). These modeling techniques generally exhibit desirable performance in precision, accuracy, or sensitivity (i.e., ability to identify a site as being in nonreference condition) but have yet to be widely incorporated into practice, perhaps because their results do not show uniform improvement over traditional RIVPACS models or because the improvements are marginal.

Back to the future: the regional reference site approach revisited.—Estimation of site-specific reference conditions through modeled interpolation circumvented one of the major concerns that Hughes (1995) had

with site-specific biological criteria—that site-specific reference expectations derived from individual sites could not be adequately extrapolated to other sites. Even though proponents of the regional reference site approach clearly recognized that “Ecoregional reference sites still require some level of habitat classification at the site scale” (Hughes 1995, p. 35), there was concern that each test site would have to be matched to ≥ 3 appropriate reference sites (i.e., replication based on matching specific site characteristics, such as catchment size or channel type), and the cost of sampling 150 to 300 reference sites each year (assuming a management agency assessed 50–100 sites/y) would be exorbitant. On the other hand, Hughes et al. (1994, p. 138) earlier had acknowledged that “Use of regional reference lakes or stream reaches (Hughes et al. 1986) to develop water resources criteria is hindered by ecoregion heterogeneity.” Hughes et al. (1994, p. 138) further stated that “...sites with substantial natural differences...in the same ecoregion should not have the same set of reference sites.”, and “If (subregionalization is impossible), reference sites for each type of natural gradient, substrate, or water quality should be selected.”

The fact that comparisons of biological assessment methods have been confounded by mixing the approach to predicting reference conditions with the type of index used (e.g., **Reynoldson et al. 1997**, Hawkins et al. 2010) is perhaps the main reason that consensus did not emerge from the index wars of the mid 1990s and early 2000s regarding how best to conduct ecological assessments (e.g., **Gerritsen 1995**, **Norris 1995**, Karr and Chu 2000, Norris and Hawkins 2000). Recent studies have demonstrated that the modeling approach is effective when applied to the development of MMIs (e.g., Oberdorff et al. 2001 [Fig. 1], 2002 [Fig. 1], Baker et al. 2005, Pont et al. 2006, **Cao et al. 2007**), and other investigators have started to evaluate the independent effects of prediction method and type of ecological index on the performance of ecological assessments. For example, Mazor et al. (2006) showed that RIVPACS or BEAST (i.e., Reynoldson et al. 1995) modeling improved assessment sensitivity, and both Hawkins (Hawkins 2006, Hawkins et al. 2010) and Davy-Bowker et al. (2006) have shown that modeling can improve the performance of different types of indices. Aroviita et al. (2009) recently showed that the performance of O/E_{taxa} indices derived from traditional RIVPACS-type models and a priori typologies were similar when indices were based on reference sites that spanned a limited geographic range, but that modeled indices performed better than typologies when geographic range increased.

Does the taxonomic group matter?—An implicit assumption regarding the general utility of any type of a priori classification in predicting reference-state biota is that the classification will be applicable to all taxonomic groups. Differential response of biota in different taxonomic groups to natural environmental gradients would require development of separate taxon-specific classifications to assess the condition of >1 taxonomic group. The degree to which different taxonomic groups show concordant patterns of variation across natural environmental gradients should provide insight regarding this assumption. Several studies have now addressed this issue (e.g., Paavola et al. 2003, 2006, Grenouillet et al. 2008, Infante et al. 2009, Virtanen et al. 2009). In general, these studies show that concordance is usually weak, although sometimes statistically significant. Fish, invertebrate, and plant taxa tend to respond most strongly to different natural environmental gradients and to different human-caused alterations to those environments (e.g., **Carlisle et al. 2008**, **Mykrä et al. 2008a**). The lack of strong concordance among taxa in their responses to both natural and human-influenced environmental gradients seriously constrains the application of regionalizations and typologies as a general tool in ecological assessments.

The issue of variable reference-site quality.—One of the most vexing issues that continue to complicate use of reference site data for predicting ecological benchmarks at test sites is that no sites are truly pristine any longer, and the degree of ecological alteration present at the least-disturbed sites can vary markedly both between and within regions (Stoddard et al. 2006, **Herlihy et al. 2008**). The consequence of using reference sites of variable quality is that assessed sites probably will be held to different standards. Thus, direct comparisons among sites is problematic at best and potentially difficult to justify to the water-policy regulated community (e.g., industry, agriculture). The issue is also problematic from a resource conservation perspective because as conditions progressively degrade over time, society's expectations regarding what is natural or acceptable can change. This shifting baseline syndrome has attracted considerable attention with respect to the conservation and management of marine resources (Pauly 1995, Dayton et al. 1998, Greenstein et al. 1998, Pinnegar and Engelhard 2008) and has similar implications for establishing expectations for freshwater ecosystems. The recent National Wadeable Stream Assessment (WSA) in the USA provides a clear illustration of how problematic this issue is (**Herlihy et al. 2008**). Furthermore, human-caused climate change will exacerbate the shifting baseline problem because even the most pristine

ecosystems that currently exist will be affected. Even if we have already sampled many pristine ecosystems, climate change will confound our ability to assess the natural variability of these systems and, thus, will compromise our ability to specify accurately the true ranges of natural variability when conducting ecological assessments.

Theoretically, the reference condition of water bodies can be estimated by various types of statistical hindcasting. Such techniques are fundamental to the reconstruction of lake environmental conditions from fossil biotic assemblages (e.g., Birks et al. 1990 [Fig. 1], Birks 1998, Simpson et al. 2005). The use of these techniques in lakes is aided by the presence of data-rich sediment cores that can be dated and that allow either quantitative or qualitative identification of the assemblage structure present prior to anthropogenic alterations. The structure of these prealteration assemblages provides a direct estimate of the biotic reference condition for individual lakes, and the use of transfer functions (reviewed by Birks 1998) that relate biotic structure to environmental conditions across modern lakes (i.e., surface sediments) can be used to infer the prealteration water quality (e.g., pH, P concentrations, water temperature) at individual lakes.

Lack of fossil records generally precludes routine use of paleolimnological methods in flowing-water ecosystems (but see Thoms et al. 1999, Gell et al. 2005), but attempts have been made to hindcast reference conditions by adjusting statistically for the relationships observed between biota and landuse alteration and related stressors. In the WSA, an attempt was made to hindcast reference conditions to more natural and, hence, comparable benchmarks by regressing MMI values observed at reference sites against the degree of alteration occurring across those sites. In this case, Herlihy et al. (2008) used a synthetic stressor gradient derived from a principal component analysis to quantify the joint variation in several stressor variables. This attempt to hindcast appeared to be only partially successful in generating benchmarks of equivalent quality because streams in the heavily altered central portions of the USA were estimated to be in better ecological condition than were streams in the less-altered Eastern Highlands.

We found 5 other papers (Dodds and Oakes 2004, Baker et al. 2005, Kilgour and Stanfield 2006, Robertson et al. 2006, and Angradi et al. 2009) that describe conceptually similar approaches to hindcasting but that used different analytical techniques to estimate reference conditions. Dodds and Oakes (2004) used analysis of covariance with ecoregions as categories and % anthropogenic land use as

covariates to predict background (reference) nutrient concentrations. Baker et al. (2005) used stepwise multiple linear regression (MLR) to predict fish, water-quality, and stream-habitat metrics from a suite of natural and anthropogenic catchment variables and then adjusted metric values to 0 levels of catchment alteration. Kilgour and Stanfield (2006) also used stepwise MLR of biotic and stream habitat variables on landuse variables to estimate reference conditions at levels of 0 land use. Robertson et al. (2006) used spatial regression-tree analyses to delineate zones of similar reference concentrations of P and suspended sediment. Angradi et al. (2009) used quantile regression to hindcast fish MMI values to 0 stress levels. In all of these examples, the authors acknowledged several potential problems in applying such models. Perhaps the most problematic issue is that of extrapolating beyond the range of calibration data, a “damned if you do, damned if you don’t” situation. By extrapolating beyond the range of the calibration data, we risk serious prediction errors; by not extrapolating, we have no estimate of what historical reference conditions were like.

Chessman and Royal (2004) explored a different approach to the variable-quality reference site dilemma. These authors built on the environmental filters concept (Poff 1997), which assumes that key environmental features screen potential colonists based on the ecological traits (tolerances) possessed by each taxon. If the original or desired environmental reference conditions at a test site can be estimated, it should be possible to predict the set of taxa that should occur at a site based on knowledge of their environmental tolerances. A comparison of the observed taxa with the predicted taxa provides an assessment measure. This approach is conceptually appealing because it is grounded in ecological first principles and, therefore, should be widely applicable and transferable. However, that it is a first principles approach also might be its chief limitation because its performance will depend critically on how well we understand the adaptive traits of individual species (e.g., Lytle and Poff 2004, Poff et al. 2006) and how well we can describe environmental filters in a way relevant to how multiple species perceive them. This approach has yet to be tested thoroughly, but it might be the only approach with the potential to overcome the issues associated with statistical extrapolation described above. Other research has shown that ecological indices based on functional ecological traits can discriminate between human-stressed and reference sites (Gayraud et al. 2003) and that trait structure might be reasonably stable across reference sites and, thus, potentially much easier to quantify and predict

than taxonomic composition (e.g., Bady et al. 2005, Statzner et al. 2005). However, Statzner et al. (2005) acknowledged that ecologists have not developed a strong predictive theory regarding the ecological reasons for the spatial and temporal variation in trait structure that has been observed among sites (see Dolédec and Statzner 2010).

Predicting Physical-Chemical Reference Conditions

The accuracy and precision with which we can establish an ecological benchmark is, in part, determined by how well we characterize and predict the natural physical and chemical conditions expected at different sites within diverse landscapes. Many of these natural conditions at a site can be affected by anthropogenic activity, so we cannot use field measurements to predict the reference-condition biota. The use of coarse surrogates, e.g., elevation for temperature, allows us to adjust predictions of biota for differences in naturally occurring environmental gradients among sites to some degree. However, such surrogates are often imprecise predictors, at best, of the environmental factors that most strongly influence the abundance and distributions of biota (e.g., local temperature, substrate, flow, and water chemistry). If we can estimate the physical-chemical reference state in terms of specific physical-chemical attributes, we should be able to predict more accurately the reference-state biota expected at a site. For example, we have developed RIVPACS-type models to predict stream biota that use the output of stream temperature models as a predictor. These RIVPACS models more accurately predict the composition of stream invertebrates assemblages than do models that use elevation, latitude, and catchment area as surrogates for stream temperature (RAH and CPH, unpublished data).

Considerable potential now exists for stream ecologists to take advantage of research conducted across various subdisciplines of the watershed sciences. Researchers studying stream temperature, hydrology, water chemistry, and channel geomorphology and substrate have made significant advances in developing methods that might be useful in predicting the physical-chemical conditions expected under reference conditions. In the following 4 sections, we describe these advances and discuss how they are being, or might be, used in ecological assessments.

Predicting thermal reference conditions

Temperature strongly influences the distributions of ectotherms (reviews by Allan 1995, Begon et al. 1996) through its effect on metabolic and growth rates

(Newell and Minshall 1978, Merritt et al. 1982), phenology (Sweeney and Vannote 1981), and fecundity (Vannote and Sweeney 1980). Therefore, accurate and precise predictions of the thermal conditions that characterize reference sites should lead to better ecological assessments.

Stream temperature researchers have approached temperature prediction mainly with 2 of the approaches used by stream ecologists to predict overall reference conditions: single-site and site-specific predictions (Table 2). We found no examples in the literature that used regionalization to characterize the thermal environment of streams. Site-specific predictions can be further divided into 2 modeling approaches and 2 spatial domains: 1) physical single-site, 2) physical multisite, 3) empirical single-site, and 4) empirical multisite. Physical single-site models typically are based on energy balance equations, whereas empirical single-site models most often relate water temperature to air temperature at a nearby temperature station. Physical multisite models typically are designed to predict temperature at many unmeasured locations within a single catchment and require measurement and parameterization of several local environmental factors. Empirical multisite models typically relate stream temperature to landscape features known to affect or covary with stream temperature, e.g., latitude, drainage area, or elevation.

Typologies.—The typological approach has not been used extensively to predict stream temperatures (Table 2). Snelder and Biggs (2002) proposed a hierarchical classification of New Zealand streams based on their climatic variability (air temperature and precipitation) and source of flow. Their classification has been tested for its ecological (Snelder et al. 2004a, b) and hydrologic (Snelder et al. 2005) predictive power, but has not been tested for how well it predicts stream temperature. Nelitz et al. (2007) used classification and regression-tree analysis to classify 104 streams in British Columbia, Canada, into 6 summer maximum weekly average temperature (MWAT) classes based on catchment size, catchment elevation, and air temperature. For each class, they used Bayesian regression to predict the probable response of MWAT to forest management practices, such as road building. Another typological approach to characterize stream temperatures is to classify streams based on the thermal preferences of the biological assemblage found within them (e.g., Lyons 1996, Wehrly et al. 2003), i.e., the fauna serve as surrogates for temperature. However, this latter approach has no utility in predicting reference-condition temperature for ecological assessment purposes because of its obvious circularity.

Site-specific models.—Site-specific physical stream temperature models emerged in the 1960s (Edinger et al. 1968, Brown 1969) largely in response to the need to assess the effects of forest and catchment management practices on aquatic ecosystems (Brown 1970). This research has produced a significant body of literature (Table 2) and some software packages (e.g., SNTMP by Theurer et al. 1984, CEQUEAU by Morin et al. 1987). The physical-based model of Theurer et al. (1984) was designed to predict both spatial and temporal variation in stream temperature within an entire stream network. Physical-based models have contributed significantly to our understanding of the processes that affect stream temperature (reviewed by Caissie 2006, Webb et al. 2008) but have not been used much in ecological assessments because they require detailed measurements of numerous fine-scale environmental variables, such as local wind speed, relative humidity, and ground reflectivity. However, Allen et al. (2007) developed a model (BASINTEMP) based on a simplified characterization of the main physical factors that influence stream heat budgets (solar radiation input, vegetation cover, groundwater input) that produces site-specific predictions within a geographical information system (GIS) framework. BASINTEMP predictions are constrained to locations within a single stream network, but this model was an important step toward development of physically based models that could be applied to larger regions.

Initial development of single-site, empirical models also occurred in the 1960s. The first models were based on simple sinusoidal functions fit to annual seasonal variation in monthly mean stream temperatures (Ward 1963). Later, linear and nonlinear regression models were developed that related stream temperature to air temperature recorded at a nearby weather station to create temporally continuous predictions of water temperature (e.g., Cluis 1972, Mohseni et al. 1998). These latter models have been used to examine the potential effects of climate warming on specific streams (Mohseni et al. 2003).

Less emphasis has been placed on developing models that predict how temperatures vary across streams (Table 2), and many of these models were developed primarily in support of ecological analyses. Miyake and Takeuchi (1951) regressed monthly mean temperature from 20 rivers in Japan against air temperature, and their model appears to be the first example of an empirical model capable of predicting at multiple sites. Over the last 30 y, several authors have used linear regression to model various measures of stream temperature (i.e., total annual degree days, daily, monthly, or annual means) as functions of

latitude (Vannote and Sweeney 1980), air temperature (Walker and Lawson 1977, Smith 1981, Stefan and Preudhomme 1993, Eaton and Scheller 1996, Pilgrim et al. 1998), or elevation (Rahel and Nibbelink 1999). Still others have used fish distributions to determine elevation isoclines of fish occurrence, which are assumed to be predictive of stream temperature (Meisner 1990, Rieman and McIntyre 1995). In addition to latitude, air temperature, and elevation, more recent research has incorporated local and catchment factors into models (e.g., catchment area, catchment slope, soil characteristics, landcover) (Jones et al. 2006, Wehrly et al. 2006).

Statistical approaches other than linear regression also have been evaluated. Gardner et al. (2003) used a geostatistical technique (kriging) to interpolate temperatures between measured sites within a catchment. Guillemette et al. (2009) used canonical correlation analysis to plot 20 stations in environmental space (catchment and site physical attributes). Once stations were projected into environmental space, the monthly maximum stream temperatures were interpolated between stations to predict temperature at unmeasured sites based on their physical attributes. Wehrly et al. (2009) compared 3 statistical techniques (MLR, generalized additive models, and linear mixed models) to predict July mean stream temperatures from catchment and site characteristics and found little difference between methods. The root mean square error reported for the 3 models was 2.0 to 2.6°C, which are comparable to errors reported in other studies (e.g., Smith 1981, Pilgrim et al. 1998). Ordinary kriging also was tested but did not perform as well as the other techniques (RMSE = 2.95°C; Wehrly et al. 2009).

Future research needs: challenges and opportunities.—Because temperature is so critical to stream biota, it is imperative that we accurately describe the thermal conditions that characterize reference streams. Many of the empirical multisite models we reviewed were developed within a single catchment, a small region, or in terrain, such as the central USA, that is topographically homogeneous. Our own analyses relating stream temperatures to GIS-derived catchment and local predictors show that complex topography, like that found in many parts of the western USA, can contribute greatly to the heterogeneity of local environmental conditions and, therefore, to errors associated with temperature prediction (RAH and CPH, unpublished data). Physical models have greatly improved our understanding of how local environmental conditions control stream temperature, and thoughtful inclusion of those factors into multisite, spatially explicit models should improve their

performance and allow large-scale spatial prediction of stream temperature as functions of both natural and anthropogenic influences.

Predicting hydrologic reference conditions

Hydrologic regime is another important aspect of the habitat template that influences biotic assemblages. Thus, improving characterization of hydrologic reference conditions expected at sites should improve ecological assessments. Stream flow can be described in many ways, but magnitude, timing, and variability appear to be particularly important to biota (e.g., Gustard 1984, Poff et al. 1997). Hydrologists were motivated initially by a need for landscape-scale predictions of flood risk, and among scientists in the 4 subdisciplines reviewed here, hydrologists have applied the widest variety of techniques to make predictions. However, empirical models and simple physically based rainfall-runoff models with empirically estimated parameters have become the prevailing approaches to making site-specific predictions of flow because of a lack of flow data and the multiscale complexity inherent in catchment dynamics. In addition, modeling has expanded from predicting one aspect of flow to predicting the entire hydrograph over time as the need for predictions has moved from assessing flood potential to predicting available water supplies, hydropower potential, and ecological flows.

Regionalizations.—Predictions of hydrologic regime began in the 1930s and 1940s with delineations of regions with homogeneous flow characteristics, topography, and climate (Krasovskaia 1997). Our use of the term, regionalization, as the development and identification of geographic regions that are homogeneous with respect to physical-chemical attributes differs from how the term is used in the hydrology literature (i.e., development of a regionally applicable model). For consistency throughout our paper, we use the former meaning of the term. Mosley (1981) derived hydrologic regionalizations via cluster analyses of hydrologic data. Others (Singh 1971, Cunnane 1988) have predicted flow for regions with reasonably homogeneous flow characteristics by applying regionally representative flows to all catchments in those regions. The regionalization approach to flow prediction does not appear to be widely used (Table 2), but the approach is still used for prediction in some cases (e.g., Chen et al. 2006).

Typologies.—Juncker (1971) developed a global hydrologic typology to predict hydrologic regimes, but since then relatively few studies have explored this approach (Table 2). Most typologies classify

catchments by their physical characteristics via cluster analysis (Acreman and Sinclair 1986). The Ecological Limits of Hydrologic Alteration framework (ELOHA; Henriksen et al. 2006, Apse et al. 2008) uses a typology to predict flow regime. Snelder et al. (2005) developed a flow regime typology that explained significantly more variation in flow among sites than did a regionalization approach.

Site-specific predictions.—Hydrologists recognize that most landscapes are too heterogeneous for effective regional classification, so they have focused on development of site-specific prediction methods that can be applied across regions (Table 2). Interpolation or mathematical modeling is used to predict flow at ungauged sites from data collected at gauged sites. Flow regimes of gauged sites have been interpolated to ungauged sites in 2 main ways (reviewed by Holmes et al. 2002): 1) spatial interpolation from data collected at those gauges that are closest to the ungauged site (e.g., Skoien et al. 2006) or 2) interpolation based on flow records collected at streams that are most similar in physical characteristics to the ungauged site (interpolation in environmental space; e.g., Acreman and Wiltshire 1989, Burn 1990). We found few studies that based predictions on interpolation approaches. Empirical and physical models were used more commonly, and both approaches build on a proposal by Nash (1960) to estimate unit hydrograph parameters from catchment characteristics via regression. The choice between empirical and physical models typically has depended on the level of detail required (NERC 1975). Stratifying natural variation among sites by applying a typology often improves predictions of both empirical and physical models (Tasker 1982, Burn and Boorman 1993), especially in hydroclimatically diverse regions (Sanborn and Bledsoe 2006).

Thomas and Benson (1970) built on earlier empirical models developed for different regions and regressed 71 different flow indices nation-wide on 21 physical catchment characteristics. Others expanded the use of empirical models by introducing weighted and generalized least squares regression (Stedinger and Tasker 1985), nonlinear models (Pandey and Nguyen 1999), ANNs (Nayebi et al. 2006), and Random Forests (Carlisle et al. 2009). Empirical models avoid some important problems associated with physical models, such as model structure uncertainty and difficulties in identifying unique parameters, and are considered more robust than physical models (Croke and Norton 2004).

In earlier studies, hydrologists developed process-based physical models to predict flow at ungauged sites by parameterizing models for ungauged sites

with parameters from models developed at gauged sites. This approach accounted for 49% of the hydrology publications we examined. Originally, physical models were calibrated with data from all gauged catchments in a region, and then model parameters were statistically related to catchment attributes, such as climate, topography, land use, geology, soils, or composites of these variables (Sefton and Howarth 1998). These attributes could then be measured at an ungauged site and the statistical relationships could be used to estimate the parameters for the model at the ungauged site (Benson and Matalas 1967, Ross 1970, Heerdegen and Reich 1974, Jarboe and Haan 1974, NERC 1975). Various physical models have been used; these models range from the simple lumped conceptual rainfall-runoff models IHACRES (Sefton et al. 1995) and SYMHID (Peel et al. 2000) to the more complex spatially distributed Stanford Watershed (Ross 1970), HBV (Seibert 1999), and SWAT models (Heuvelmans et al. 2006). Fernandez et al. (2000) refined this approach by simultaneously maximizing the fit of model parameters to catchment characteristics and gauge data. Physical models also can be calibrated to ungauged sites by spatially interpolating parameter values from gauged sites, but this approach has produced mixed results (Merz and Blöschl 2004, 2005, McIntyre et al. 2005, Oudin et al. 2008). Other refinements include constraining parameter space during model calibration (Gotzinger and Bardossy 2007), relating model parameters to catchment attributes using weighted or sequential regression (Kay et al. 2006), and making predictions from ensembles of models (McIntyre et al. 2005). Increasingly, site-specific predictions are made with either complex physically based models that exploit spatially extensive fine-scale data (Ivanov et al. 2004, Immerzeel and Droogers 2008) or new techniques for estimating model parameters, like ANNs (Heuvelmans et al. 2006).

Future research needs: challenges and opportunities.—Since 1960, hydrologists have developed multiple techniques that potentially could be used to predict the reference hydrologic regime for any site within a given region. Recent research has focused on how to quantify and reduce the uncertainty associated with these predictions (Sivapalan et al. 2003) and how to apply predictions to ecological assessments (Henriksen et al. 2006, Sanborn and Bledsoe 2006, Kennen et al. 2008). Sivapalan et al. (2003) established the International Association of Hydrological Sciences (IAHS) Decade on Predictions in Ungauged Basins to focus research efforts on measuring and reducing prediction uncertainties, and several researchers have begun to address this challenge (e.g., Wagener and

Wheater 2006, Zhang et al. 2008). Sanborn and Bledsoe (2006) used an empirical model to predict 84 streamflow metrics (with estimates of uncertainty) that describe flow regime at 150 ungauged sites across 3 states in the US. They also explained how these predictions could be used to: 1) improve ecological assessments by predicting flow characteristics affecting taxa distributions, 2) assess causes of impairment by quantifying flow disturbances, or 3) design a restoration that maximizes recruitment of cottonwood trees. Kennen et al. (2008) estimated parameters for a physical model from spatial data (i.e., topography, soils, impervious surface cover, precipitation, and temperature) and then predicted reference flow conditions. They compared reference flow to current flow conditions and related these differences in flow to macroinvertebrate assemblage composition. Advances in hydrologic modeling have made predictions of the hydrologic regime possible, and these predictions have great potential to improve models of biotic-environment relationships. We see no reason why these hydrologic models should not be used more frequently in the future to provide better understanding of the reference condition of streams.

Predicting geomorphic and bed sediment reference conditions

Structural habitat is a key factor that determines the spatial distribution of aquatic taxa and biotic assemblages. Here, we define structural habitat as the channel geomorphic characteristics and sediment size distributions that compose the physical environment upon and within which stream biota live (Minshall 1984). Stream geomorphic structure and sediment size are important to biota, so ecological assessments should benefit from their improved prediction at unmeasured sites. Studies of stream structural habitat often consider aspects of channel planform (e.g., straight, meandering, or braided), stream channel width, channel depth, channel slope, and streambed sediment size distributions. The physical characteristics of streams change with stream size (Leopold et al. 1964) and landscape setting (Lotspeich 1980).

Regionalizations.—We found no examples of attempts to classify whole regions with respect to either the geomorphic or sediment characteristics expected in streams (Table 2). However, some researchers have used regionalization as a first step towards a hierarchical classification of channel types (e.g., Maxwell et al. 1995).

Typologies.—Geomorphologists recognized early that streams and rivers show systematic patterns in structure that are amenable to classification (Gilbert

1877, Davis 1890). Numerous classification systems have been proposed (Table 2) based on stream size (Horton 1945, Strahler 1957), channel planform (Leopold and Wolman 1957), or the dominant geomorphic processes that form streams (Schumm 1977). More recently, Lotspeich (1980) proposed a hierarchical framework for catchments, and Frissell et al. (1986) refined this classification by adding finer nested levels of geomorphic units (catchment > stream system > valley segment > reach > pool/riffle > microhabitat). Much research in the 1990s focused on describing the structure of these units and the processes that create them, e.g., valley segments (Nanson and Croke 1992), reaches (Rosgen 1994, Montgomery and Buffington 1997), and pools and riffles (Hawkins et al. 1993). To the best of our knowledge, no typologies have been proposed that focus exclusively on differences in substrate size among reaches.

Several classification systems have been suggested as management tools for aquatic resources (e.g., Maxwell et al. 1995). These classification systems were designed to inform land managers about how different stream classes should respond to management activities (Downs 1995), were often created for a specific region (Paustian et al. 1992), and can include stream condition as part of the classification scheme (Rosgen 1996). Kondolf et al. (2003) provide a comprehensive discussion of geomorphic classification systems and their purposes and uses.

Site-specific predictions.—Physical models that treat sediment size as an independent variable to predict channel slope, sediment movement, and yield (Dade and Friend 1998) can be rearranged to predict sediment size at unmeasured locations. However, few studies have attempted to make site-specific predictions of sediment size from physical relationships (Table 2). Here, we briefly review those efforts and describe several recent attempts to develop empirical models that predict sediment characteristics for use in aquatic assessments.

Montgomery et al. (1998) suggested that Shields' equation (Shields 1936), which describes the shear stress needed to move the median sediment size (D_{50}) in a stream reach at bankfull discharge, can be rearranged to predict D_{50} from bankfull depth and channel slope. Buffington et al. (2004) used an approach similar to analysis of covariance to show that certain geomorphic river classes (Montgomery and Buffington 1997) have smaller sediment sizes than predicted from Shields' equation. This analysis was used to adjust sediment size predictions in 3 river networks in Washington, USA. Buffington et al. (2004) used drainage area to estimate bankfull depth and a

10-m-resolution digital elevation model (DEM) to derive channel slope. Kaufmann et al. (2009) rearranged the Shields' equation to estimate relative bed stability (RBS) in 101 streams in the Pacific Coastal Range, USA. These authors related RBS to a gradient of land use and found that reference-condition streams tended to be more stable at high flows. In addition, Kaufmann et al. (2009) used MLR to predict RBS values from field and GIS-derived variables ($R^2 = 0.62$). The selected predictor variables included basin lithology, a surrogate for unit stream power ($A_{ws}^{0.5}S$, where A_{ws} is basin area and S is the channel slope), and management-related factors (road density, a riparian condition index), results implying that reference-site sediment character depends largely on discharge (function of watershed area) and basin geology.

Several authors recently attempted to predict sediment size with empirical models that used GIS-derived predictors. Flores (2004) regressed sediment D_{50} estimates from stream sites in Oregon and Colorado, USA, against DEM channel slope, drainage area, elevation, 2 catchment lithology variables, 2 valley confinement variables, and mean precipitation. His model explained 84% of the variation in D_{50} , but probably suffered from some overfitting, given the sample size ($n = 39$) and number of predictors (8). Two other studies (Mugodo et al. 2006, Frappier and Eckert 2007) used a suite of GIS variables to predict various local habitat features including sediment characteristics (e.g., % sand, % cobble, % gravel). Both models suffered from large prediction errors. However, the modeled estimates of sediment size in Mugodo et al. (2006) predicted the distributions of 7 fish species as effectively as field measurements of sediment size, results implying that their modeled sediment sizes were predicted well enough to be ecologically useful.

Future research needs: challenges and opportunities.—Use of geomorphic typologies can aid in conducting inventories of aquatic resources (Brenden et al. 2007) and predicting the biotic assemblages expected at unvisited locations (Naiman 1998). Advances in GIS have allowed deployment of various typology schemes that account for a suite of habitat characteristics, including geomorphology, and that have promise for use in bioassessments (e.g., **Snelder et al. 2004b**).

Future progress in sediment size prediction will require expansion of the scale of investigations (i.e., larger sample sizes and geographic extents) and incorporation of local processes known to affect sediment size. However, incorporation of local processes at the spatial scale required for many ecological

assessments might be challenging, especially for small streams where stochastic events, such as a single tree fall within the channel, can have a disproportionately large effect on the geomorphic structure when compared with downstream reaches (e.g., Swanson et al. 1976, Keller and Swanson 1979, **Gooderham et al. 2007**). Numerous studies have examined the factors controlling downstream variation in sediment size in individual rivers (e.g., Brierley and Hickin 1985, Hoey and Bluck 1999, Rice 1999). Recently, these papers have emphasized the importance of local natural controls that disrupt the idealized trends in the downstream reduction in sediment size (Ferguson et al. 2006). Local controls can include lateral sediment contributions from hill slopes (Benda and Dunne 1997, Bigelow et al. 2007) and tributaries (Rice 1998, reviewed by Benda et al. 2004), adjustments in flow and channel competence resulting from large woody debris inputs (Massong and Montgomery 2000, Gomi et al. 2006, Kaufmann et al. 2009; see Van Sickle and Gregory 1990, Sobota et al. 2006, Wohl and Jaeger 2009 for models of large wood recruitment to streams), fine sediment retention by small mountain lakes (Arp et al. 2007), and local geology (Mureşan 2009). From a practical application context, it is important to note that each of the natural processes listed above potentially could be mapped in a GIS. In addition, these processes have received much attention recently, and significant progress has been made toward understanding why, where, and how much each factor affects local stream sediment size. Additional progress toward predicting reference sediment characteristics will require explicit incorporation of these natural processes in modeling efforts.

Predicting water-chemistry reference conditions

Water chemistry, defined as the concentrations of major cations (e.g., Ca^{2+} , Mg^{2+}), major anions (e.g., Cl^- , SO_4^{4-}), and nutrients (e.g., N, P), is important as a direct measure of water quality and as a factor structuring freshwater assemblages. For example, Egglisshaw and Morgan (1965), Sutcliffe and Carrick (1973), and Minshall and Minshall (1978) showed that streams with low ionic strength and pH had low taxon richness and abundances, and experiments done by Willoughby and Mappin (1988) showed that water chemistry restricts the distributions of some taxa. Therefore, describing water chemistry expected under reference conditions should improve our ability to describe the ecological benchmarks appropriate under different landscape settings.

In the 1980s and 1990s, most predictive modeling efforts for water chemistry focused on the effects of

acid rain on water chemistry, and $>1/2$ of the publications we reviewed addressed this one aspect. Acidification research was directed mainly toward development of physical models that were calibrated with measured water-chemistry and atmospheric inputs and that were designed to make temporal predictions at individual sites (reviewed by Reuss et al. 1986, Christophersen et al. 1993, and Breuer et al. 2008). Exceptions to this tendency include development of regional, typological, or empirical modeling approaches to predict water chemistry to aid in determining attainable ecosystem conditions or predicting taxon distributions.

Paleolimnological approaches that relate water-chemistry conditions to the distributions of diatom or chironomid taxa have been developed for lakes (e.g., Dixit et al. 1999). These models can be applied to the taxa observed in sediment cores to estimate historical water-chemistry conditions at individual sites. When coupled with information on core dates, such models can identify water-chemistry conditions at individual lakes prior to human alteration of the landscape. However, these models cannot directly estimate the water-chemistry conditions expected at unsampled lakes.

Catchment lithology is one of the major factors controlling water chemistry, so geologic data sometimes can be used as a surrogate for water chemistry when describing how ecological properties vary with water chemistry or when creating ecologically relevant regionalizations or typologies. For example, Leathwick et al. (2005) used measurements of rock P and Ca concentrations as potential predictors in models of fish distributions. Geology also is used to control for variations in water chemistry in the European Union's WFD System-A typology (Davy-Bowker et al. 2006).

Regionalizations and typologies.—Landscape classifications comprised a minority of the publications we found regarding prediction of water chemistry (Table 2), but they were important to understanding variation in water chemistry and continue to be used in both bioassessments and in the establishment of nutrient criteria (**Johnson and Host 2010**). The development of a method for predicting stream alkalinities based solely on geologic regions by Bricker and Rice (1989) represents one of the earliest uses of a regionalization to predict water chemistry, i.e., the proportion of catchments with acid neutralizing capacity $<200 \mu\text{eq/L}$ within 1% of independently measured values. Ecoregions (Omernik 1987) are the most frequently used regionalization and might be expected to partition some of the variation in water chemistry because they are partly derived from

2 primary drivers of water-chemistry variability—lithology and land use. However, the success of ecoregions in partitioning the spatial variation in water chemistry has been mixed (Harding et al. 1997, **Pan et al. 2000**, Jenerette et al. 2002, Rohm et al. 2002, **Herlihy and Sifneos 2008**). Typological approaches also have incorporated aspects of predicted water chemistry into catchment and reach classifications, e.g., Snelder and Biggs (2002) and the WFD System-A typology (Davy-Bowker et al. 2006). Both regional and typological methods have been important in recent efforts to establish nutrient reference conditions and criteria (Rohm et al. 2002, Robertson et al. 2006), although some investigators have raised doubts about the ability of a classification approach to partition natural variation sufficiently well to establish defensible nutrient criteria (e.g., Smith et al. 2003, **Herlihy and Sifneos 2008**).

Site-specific predictions.—Most water-chemistry predictive models have been developed to make predictions at single sites. The pioneering work of Johnson et al. (1969), who modeled stream water chemistry as a joint function of soil water and precipitation with empirically derived parameters, led to much of the later research that explored both empirical and physical modeling approaches. However, physical modeling quickly became the dominant approach once the Birkenes Model had been developed (Christophersen et al. 1982). This model is composed of submodels for hydrology and different ion species and produces estimates of water chemistry over time at a site. Later models followed this approach, expanded the number of ionic species modeled, and included nutrient concentrations. Papers describing physical models now constitute 50% of the literature, with MAGIC (Cosby et al. 1985), ILWAS (Goldstein et al. 1984, Gherini et al. 1985), PnET (Aber and Federer 1992), and INCA (Whitehead et al. 1998) models among those most often used. These models can make temporal predictions, but they cannot predict water chemistry at unmeasured locations or be used to assess the impact of land use because they must be calibrated with water chemistry measured at each site. Most models that potentially would be capable of site-specific predictions still require site measurements of water chemistry (Aherne et al. 2003) or model entire regions (a form of regionalization) and do not make site-specific predictions (Hornberger et al. 1989).

Stumm et al. (1983; see also Schnoor and Stumm 1986) developed simpler steady-state physical models for site-specific predictions based on mass-balance relationships, without incorporating hydrology into the model. Publications describing these types of

models represent ~12% of the literature, and AL-CHEMI (Schecher and Driscoll 1988), SSWC (Henriksen et al. 1992), and PROFILE (Warfvinge and Sverdrup 1992) models are the leading examples. Because of their simple structure and low data requirements, these models often are used to assess the sensitivity of water bodies to acid loading across regions from large data sets of site measurements, but predictions still require ≥ 1 site measurement (but see the regional version of the SSWC model in Jones et al. 1990). The PROFILE model (Warfvinge and Sverdrup 1992) is an exception and can make spatial predictions of water chemistry from mapped soil properties.

Publications describing physical models dominated the literature in the 1980s, but their limitations in making predictions at multiple unmeasured sites stimulated development of empirical approaches. For example, Billett and Cresser (1992) developed site-specific empirical models that predicted water chemistry from mapped soil chemistry (see also Small and Sutton 1986). This work was later expanded to include predictions based on multiple predictor variables (Smart et al. 1998, 2001, Cresser et al. 2000). Publications focused on site-specific empirical modeling now make up ~26% of the literature. End Member Mixing Analysis (Christophersen et al. 1990, Hooper et al. 1990) is another important empirical technique but is not practical for use at large scales because it requires more data than do physical models. Geostatistical techniques have been used to predict water chemistry in 4 studies and have produced better predictions than have empirical models (e.g., Peterson et al. 2006).

Future research needs: challenges and opportunities.—The empirical modeling approaches described above and recent developments in physical and statistical modeling should allow ecologists to estimate better the water-chemistry characteristics of reference conditions. In some cases, these models are already in use. For example, Baker et al. (2005) used empirical models to establish reference chemical conditions in the upper Midwest, USA, as part of an ecological assessment, and then compared these predictions to current measurements to show that 22 to 35% of sites were impaired. Soranno et al. (2008) used empirical models to establish reference P conditions for 374 lakes in Michigan, USA. Physical models are now being used to estimate preindustrial water chemistry (Erlandsson et al. 2008, Wade et al. 2008, Zhai et al. 2008), but these models account only for changes in atmospheric deposition. Physical models also are being combined with empirical models to allow spatially explicit predictions. Chen and Driscoll (2005) used empirical models and GIS data to

parameterize the PnET-BGC model and applied the model across the Adirondack Mountains in New York, USA. Evans et al. (2006) linked the MAGIC model with an empirical model to predict water chemistry spatially throughout a catchment and temporally from 1850 to 2050. They related these predictions to invertebrate taxon distribution models to show how invertebrate distributions changed as acidity increased. Newer statistical approaches, like ANNs, also are yielding site-specific predictions with reasonable accuracy (pH within 2% of measured values, Ehrman et al. 1996; a model of total N with $R^2 = 0.94$, Amiri and Nakane 2009). Despite these advances, the amount of uncertainty in predictions based on both physical and empirical approaches must be assessed fully, especially when the models are applied to heterogeneous terrains (Hill et al. 2002).

*Reference conditions, numeric criteria,
and water-quality standards*

Quantitative ecological assessments ultimately require that numeric criteria be established that describe either the maximum acceptable level of a water-quality element (e.g., a nutrient concentration) or the minimum acceptable value of a biotic index. In the past, both regional and typological methods have been used to establish water-quality criteria (e.g., Rohm et al. 2002, Robertson et al. 2006) and to develop biocriteria (USEPA 2002). However, water chemistry, temperature, sediment, and flows can vary markedly within ecoregions (e.g., Lewis et al. 1999, **Dodds and Welch 2000**), hence a one-size-fits-all approach to establishing numeric criteria almost certainly would produce water-quality or ecological standards that either are not realistically attainable given natural conditions (e.g., Smith et al. 2003) or are not protective enough (e.g., Dodds and Oakes 2004). Applying site-specific models instead of regional or typological methods when developing numeric criteria should allow establishment of water-quality and biological criteria that better match the true potential of individual sites.

Physical-chemical predictions and ecological assessments

Advances in the watershed sciences have greatly increased our ability to predict the physical-chemical habitat at a site, and these advances should be exploited by the ecological assessment community. We found examples of successful predictions of the physical-chemical attributes expected under reference conditions in all 4 watershed sciences subdisciplines that we reviewed. Linking these predictions with the biota observed at reference sites should allow us to

establish more robust biota–environment relationships that should, in turn, improve precision and reduce bias of ecological assessments. For example, predicted water temperature could be used instead of surrogates of temperature, such as elevation, in RIVPACS-type predictive models to estimate E. Our modeling of taxon distributions in Colorado, USA, streams showed that predicted reference-state stream temperatures led to more accurate predictions of distributions of stream biota than did temperature surrogates (RAH and CPH, unpublished data). In addition, use of predicted site-specific physical-chemical attributes in biota–environment models should aid in interpreting the causal relationships embedded in these empirical models, something that is difficult to do when surrogate predictors are used. The use of predictors that have a likely causal link to the distribution of biota also should aid in the diagnosis of at least some causes of biological impairment, i.e., those landscape or waterway alterations that affect stream temperature, hydrology, geomorphology, and water chemistry. These benefits alone might justify the use of modeled physical-chemical stream characteristics when predicting reference-state biota, even if they do not significantly improve the precision or bias of ecological assessments.

The 4 watershed sciences subdisciplines that we examined are not equally mature. For example, the ability to predict hydrologic regimes is further developed than the ability to predict in the other 3 disciplines. Hydrologic regime papers comprised almost ½ of the papers that describe models capable of site-specific predictions of reference conditions (Table 2). The reliance of most physical models developed for temperature and water chemistry on site-specific parameterization currently limits their use in establishing expected site-specific reference conditions in 2 ways. The need for multiple site measurements makes application to large populations of streams prohibitively expensive, but more importantly, these measurements (e.g., soil water chemistry or channel width–depth ratios) often are affected by anthropogenic stressors, a trait that makes any prediction of reference conditions suspect. Scientists working in other subdisciplines should consider exploring methods used by hydrologists to develop physical models that can be applied regionally without relying on site measurements. We predict that the accuracy and precision of biota–environment models, and hence ecological assessments, will improve as site-specific predictions for each of these physical-chemical attributes are more widely applied.

Regional and typological classifications also have been used to partition some of the natural variation in

physical-chemical conditions (Table 2). However, the collapsing of continuous gradients into classes that is inherent in these approaches might not be necessary given the state of development of both empirical and physical models. Model predictions in all 4 subdisciplines also have been used to help explain distributional patterns in biota, but only hydrologic predictions have been applied explicitly in an ecological assessment context. We think that ecological assessments would benefit from increased use of temperature, hydrologic regime, water chemistry, and substrate predictive models to predict reference conditions. Applications of these models should be tested explicitly to determine how much they improve ecological predictions. As these modeling approaches continue to evolve to include measures of prediction uncertainty, ecologists also should develop methods to assess and account for the propagation of these errors when these models are applied to ecological assessments.

Concluding Remarks and Recommendations

Our review showed that the science supporting ecological assessments has matured greatly over the past 25 y (*sensu* **Creutzburg and Hawkins 2008**). This maturation occurred as a consequence of: 1) a large and active bioassessment research community within the North American Benthological Society and elsewhere and 2) robust testing and refinement of those concepts and methods most critical to assessing the status of ecological resources. Research in the area of ecological assessment has been so active that we were unable to cover all of the contributions that have been made regarding the many factors that influence the estimation of reference conditions. Considering the ever-growing threats to the planet's ecological resources, we will need to refine further our ability to conduct accurate and precise ecological assessments of freshwater ecosystems. In doing so, we expect to see researchers and managers who work on other types of ecosystems take better advantage of the wealth of knowledge that the freshwater bioassessment community has produced. We also expect to see more cross-disciplinary research between ecological and physical scientists interested in predicting reference conditions. In all cases, future progress will benefit greatly if research is collaborative in nature and actively questioning in spirit.

From our review, we think sufficient information now exists to support the following recommendations: 1) We should move forcefully toward adopting methods that make site-specific prediction the stan-

dard approach to estimating reference conditions. Regionalizations, typologies, and other a priori classification schemes can produce coarse estimates of reference conditions, but we do not think they have the potential to improve the accuracy and precision needed by resource managers to set appropriate numerical criteria or to detect ecologically meaningful deviation from those criteria. Future efforts should be directed toward predictive modeling approaches that can deliver these improvements. 2) We need to understand and characterize site-specific aspects of natural variability in both the ecological and physical-chemical attributes of freshwater ecosystems and to distinguish this source of variability from sampling and prediction error. 3) We need to focus specifically on how to minimize or account for systematic prediction bias in ecological assessments. In our view, the consequences of ignoring or not accounting for prediction bias could result in more serious assessment errors than do problems associated with poor precision. 4) We need to make research designed to estimate and adjust for differences in reference-site quality a priority. The effect of variation in reference-site quality on both within- and cross-region comparisons of ecological condition remains a serious and unresolved issue. Variation in reference-site quality is probably an equally, if not more, serious problem affecting comparisons of ecological assessments than is the use of either different methods of predicting the reference condition or different types of indices.

We must continue to bring better order to the science that supports ecological assessments. Clearly, some aspects of both ecological systems and the way humans develop and use ecological knowledge can appear to be chaotic, but research published in *J-NABS* and elsewhere has contributed greatly to a process that brings needed order to both the applied science of ecological assessment and the basic science on which sound ecological assessments depend. We congratulate the authors, editors, and manuscript reviewers who, through 25 y of dedicated work, have made *J-NABS* one of the world's leading journals in the field of ecological assessment.

Acknowledgements

We thank Nora Burbank, Brian Creutzburg, and Ellen Wakeley of the Utah State University Western Monitoring Center and Aquatic Ecology Lab for constructive comments on earlier drafts of the manuscript. Comments by 2 anonymous referees and Associate Editor Bruce Chessman helped hone our thinking and presentation. CPH extends a special thanks to John Van Sickle for generously sharing his

insights regarding modeling and statistical hypothesis testing in ecological assessments. Our own research on the topic of reference conditions and ecological benchmarks was supported by grants R-828637-01 and R-830594-01 from the National Center for Environmental Research (NCER) Science to Achieve Results (STAR) Program of the US EPA. We thank the *J-NABS* editors for the opportunity to discuss the evolution of this important topic.

Literature Cited

- ABER, J. D., AND C. A. FEDERER. 1992. A generalized, lumped-parameter model of photosynthesis, evapotranspiration and net primary production in temperate and boreal forest ecosystems. *Oecologia* (Berlin) 92:463–474.
- ACREMAN, M. C., AND C. D. SINCLAIR. 1986. Classification of drainage basins according to their physical characteristics: an application for flood frequency analysis in Scotland. *Journal of Hydrology* 84:365–380.
- ACREMAN, M., AND S. WILTSHIRE. 1989. The regions are dead; long live the regions. Methods of identifying and dispensing with regions for flood frequency analysis. Pages 175–188 in *FRIENDS Conference*. International Association of Hydrological Sciences, Bolkesjo, Norway.
- AHERNE, J., P. J. DILLON, AND B. J. COSBY. 2003. Acidification and recovery of aquatic ecosystems in south central Ontario, Canada: regional application of the MAGIC model. *Hydrology and Earth System Sciences* 7:561–573.
- ALLAN, J. D. 1995. *Stream ecology: structure and function of running waters*. Chapman and Hall, Dordrecht, The Netherlands.
- ALLEN, D., W. E. DIETRICH, P. F. BAKER, F. LIGON, AND B. ORR. 2007. Development of a mechanistically-based, basin-scale stream temperature model: applications to cumulative effects modeling. Pages 11–24 in R. B. Standiford, G. A. Giusti, Y. Valachovic, W. J. Zielinski, and M. J. Furniss (editors). *Proceedings of the Redwood Region Forest Science Symposium: What Does the Future Hold?* Forest Service, US Department of Agriculture, Albany, California.
- AMIRI, B. J., AND K. NAKANE. 2009. Comparative prediction of stream water total nitrogen from land cover using Artificial Neural Network and multiple linear regression approaches. *Polish Journal of Environmental Studies* 18:151–160.
- ANGRADI, T., M. PEARSON, T. JICHA, D. TAYLOR, D. BOLGRIEN, M. MOFFETT, K. BLOCKSOM, AND B. HILL. 2009. Using stressor gradients to determine reference expectations for great river fish assemblages. *Ecological Indicators* 9:748–764.
- APSE, C., M. DEPHILIP, J. ZIMMERMAN, AND M. P. SMITH. 2008. Developing instream flow criteria to support ecologically sustainable water resource planning and management. The Nature Conservancy, Harrisburg, Pennsylvania. (Available from: http://www.depweb.state.pa.us/watershedmgmt/lib/watershedmgmt/water_allocation/pa_instream_flow_report_tnc_growing_greener_final.pdf)
- AROVIITA, J., E. KOSKENNIEMI, J. KOTANEN, AND H. HÄMÄLÄINEN. 2008. A priori typology-based prediction of benthic macroinvertebrate fauna for ecological classification of rivers. *Environmental Management* 42:894–906.
- AROVIITA, J., H. MYKRÄ, T. MUOTKA, AND H. HÄMÄLÄINEN. 2009. Influence of geographical extent on typology- and model-based assessments of taxonomic completeness of river macroinvertebrates. *Freshwater Biology* 54:1774–1787.
- ARP, C. D., J. C. SCHMIDT, M. A. BAKER, AND A. K. MYERS. 2007. Stream geomorphology in a mountain lake district: hydraulic geometry, sediment sources and sinks, and downstream lake effects. *Earth Surface Processes and Landforms* 32:525–543.
- BADY, P., S. DOLÉDEC, C. FESL, S. GAYRAUD, M. BACCHI, AND F. SCHÖLL. 2005. Use of invertebrate traits for the biomonitoring of European large rivers: the effects of sampling effort on genus richness and functional diversity. *Freshwater Ecology* 50: 159–173.
- BAKER, E. A., K. E. WEHRLY, P. W. SEELBACH, L. WANG, M. J. WILEY, AND T. SIMON. 2005. A multimetric assessment of stream condition in the Northern Lakes and Forests ecoregion using spatially explicit statistical modeling and regional normalization. *Transactions of the American Fisheries Society* 134: 697–710.
- BALESTRINI, R., M. CAZZOLA, AND A. BUFFAGNI. 2004. Characterising hydromorphological features of selected Italian rivers: a comparative application of environmental indices. *Hydrobiologia* 516:365–379.
- BARBOUR, M. T., J. GERRITSEN, G. E. GRIFFITH, R. FRYDENBORG, E. MCCARRON, J. S. WHITE, AND M. L. BASTIAN. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* 15:185–211.
- BARBOUR, M. T., J. GERRITSEN, B. D. SNYDER, AND J. B. STRIBLING. 1999. *Rapid bioassessment protocols for use in streams and rivers: periphyton, benthic macroinvertebrates, and fish*. 2nd edition. EPA 841-B-99-002. Office of Water, US Environmental Protection Agency, Washington, DC.
- BÊCHE, L. A., AND V. H. RESH. 2007. Short-term climatic trends affect the temporal variability of macroinvertebrates in California 'Mediterranean' streams. *Freshwater Biology* 52:2317–2339.
- BEGON, M., J. L. HARPER, AND C. R. TOWNSEND. 1996. *Ecology: individuals, populations and communities*. 3rd edition. Blackwell Science Ltd, Oxford, UK.
- BENDA, L., K. ANDRAS, D. MILLER, AND P. BIGELOW. 2004. Confluence effects in rivers: interactions of basin scale, network geometry, and disturbance regimes. *Water Resources Research* 40: W05402. doi:10.1029/2003WR002583.
- BENDA, L., AND T. DUNNE. 1997. Stochastic forcing of sediment supply to channel networks from landsliding and debris flow. *Water Resources Research* 33:2849–2863.
- BENSON, M., AND N. MATALAS. 1967. Synthetic hydrology based on regional statistical parameters. *Water Resources Research* 3: 931–935.
- BIGELOW, P. E., L. E. BENDA, D. J. MILLER, AND K. M. BURNETT. 2007. On debris flows, river networks, and the spatial structure of channel morphology. *Forest Science* 53:220–238.
- BILLETT, M. F., AND M. S. CRESSER. 1992. Predicting stream-water quality using catchment and soil chemical characteristics. *Environmental Pollution* 77:263–268.
- BIRKS, H. J. B., D. G. FREY, AND E. S. DEEVEY. 1998. Review 1: Numerical tools in palaeolimnology—progress, potentialities, and problems. *Journal of Paleolimnology* 20:307–332.
- BIRKS, H., J. LINE, S. JUGGINS, A. STEVENSON, AND C. TER BRAAK. 1990. Diatoms and pH reconstruction. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 327: 263–278.
- BOWMAN, M. F., AND K. M. SOMERS. 2006. Evaluating a novel Test Site Analysis (TSA) bioassessment approach. *Journal of the North American Benthological Society* 25:712–727.
- BRADLEY, D. C., AND S. J. ORMEROD. 2001. Community persistence among stream invertebrates tracks the North Atlantic Oscillation. *Journal of Animal Ecology* 70:987–996.

- BRENDEN, T. O., L. Z. WANG, R. D. CLARK, P. W. SEELBACH, AND J. LYONS. 2007. Comparison between model-predicted and field-measured stream habitat features for evaluating fish assemblage-habitat relationships. *Transactions of the American Fisheries Society* 136:580–592.
- BREUER, L., K. B. VACHE, S. JULICH, AND H. G. FREDE. 2008. Current concepts in nitrogen dynamics for mesoscale catchments. *Hydrological Sciences Journal – Journal des Sciences Hydrologiques* 53:1059–1074.
- BRICKER, O. P., AND K. C. RICE. 1989. Acidic deposition to streams: a geology-based method predicts their sensitivity. *Environmental Science and Technology* 23:379–385.
- BRIERLEY, G. J., AND K. FRYIRS. 2000. River styles, a geomorphic approach to catchment characterization: implications for river rehabilitation in Bega catchment, New South Wales, Australia. *Environmental Management* 25:661–679.
- BRIERLEY, G. J., AND E. J. HICKIN. 1985. The downstream gradation of particle sizes in the Squamish River, British Columbia. *Earth Surface Processes and Landforms* 10:597–606.
- BRINSON, M. M. 1993. A hydrogeomorphic classification for wetlands. Wetlands Research Program Technical Report WRP-DE-4. US Army Corps of Engineers, Greenville, North Carolina.
- BRINSON, M. M., AND R. RHEINHARDT. 1996. The role of reference wetlands in functional assessment and mitigation. *Ecological Applications* 6:69–76.
- BROWN, G. W. 1969. Predicting temperatures of small streams. *Water Resources Research* 5:68–75.
- BROWN, G. W. 1970. Predicting the effect of clearcutting on stream temperature. *Journal of Soil and Water Conservation* 25:11–13.
- BUFFINGTON, J. M., D. R. MONTGOMERY, AND H. M. GREENBERG. 2004. Basin-scale availability of salmonid spawning gravel as influenced by channel type and hydraulic roughness in mountain catchments. *Canadian Journal of Fisheries and Aquatic Sciences* 61:2085–2096.
- BURN, D. H. 1990. Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resources Research* 26:2257–2265.
- BURN, D. H., AND D. B. BOORMAN. 1993. Estimation of hydrological parameters at ungauged catchments. *Journal of Hydrology* 143:429–454.
- CAISSIE, D. 2006. The thermal regime of rivers: a review. *Freshwater Biology* 51:1389–1406.
- CAO, Y., C. P. HAWKINS, J. OLSON, AND M. A. KOSTERMAN. 2007. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators for Idaho streams. *Journal of the North American Benthological Society* 26:566–585.
- CARLISLE, D. M., J. FALCONE, D. M. WOLOCK, M. R. MEADOR, AND R. H. NORRIS. 2009. Predicting the natural flow regime: models for assessing hydrological alteration in streams. *River Research and Applications*. doi: 10.1002/rra.1247.
- CARLISLE, D. M., C. P. HAWKINS, M. R. MEADOR, M. POTAPOVA, AND J. FALCONE. 2008. Biological assessments of Appalachian streams based on predictive models for fish, macroinvertebrate, and diatom assemblages. *Journal of the North American Benthological Society* 27:16–37.
- CHEN, L. M., AND C. T. DRISCOLL. 2005. Regional assessment of the response of the acid-base status of lake watersheds in the Adirondack region of New York to changes in atmospheric deposition using PnET-BGC. *Environmental Science and Technology* 39:787–794.
- CHEN, Y. D., G. HUANG, Q. X. SHAO, AND C. Y. XU. 2006. Regional analysis of low flow using L-moments for Dongjiang basin, South China. *Hydrological Sciences Journal – Journal des Sciences Hydrologiques* 51:1051–1064.
- CHESSMAN, B. C., K. A. FRYIRS, AND G. J. BRIERLEY. 2006. Linking geomorphic character, behaviour and condition to fluvial biodiversity: implications for river management. *Aquatic Conservation – Marine and Freshwater Ecosystems* 16:267–288.
- CHESSMAN, B., M. MUSCHAL, AND M. ROYAL. 2008. Comparing apples with apples: use of limiting environmental differences to match reference and stressor-exposure sites for bioassessment of streams. *River Research and Applications* 24:103–117.
- CHESSMAN, B. C., AND M. J. ROYAL. 2004. Bioassessment without reference sites: use of environmental filters to predict natural assemblages of river macroinvertebrates. *Journal of the North American Benthological Society* 23:599–615.
- CHRISTOPHERSEN, N., C. NEAL, AND R. P. HOOPER. 1993. Modeling the hydrochemistry of catchments: a challenge for the scientific method. *Journal of Hydrology* 152:1–12.
- CHRISTOPHERSEN, N., C. NEAL, R. P. HOOPER, R. D. VOGT, AND S. ANDERSEN. 1990. Modeling streamwater chemistry as a mixture of soilwater end-members: a step towards 2nd-generation acidification models. *Journal of Hydrology* 116:307–320.
- CHRISTOPHERSEN, N., H. M. SEIP, AND R. F. WRIGHT. 1982. A model for streamwater chemistry at Birkenes, Norway. *Water Resources Research* 18:977–996.
- CLARKE, R. 2000. Uncertainty in estimates of river quality based on RIVPACS. Pages 39–54 in J. F. Wright, D. W. Sutcliffe, and M. T. Furze (editors). *Assessing the biological quality of fresh waters: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, Cumbria, UK.
- CLARKE, R. T., M. T. FURZE, 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.
- CLARKE, R. T., AND D. HERING. 2006. Errors and uncertainty in bioassessment methods – major results and conclusions from the STAR project and their application using STARBUGS. *Hydrobiologia* 566:433–439.
- CLARKE, R. T., J. F. WRIGHT, AND M. T. FURZE. 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. *Ecological Modelling* 160:219–233.
- CLUIS, D. A. 1972. Relationship between stream water temperature and ambient air temperature. *Nordic Hydrology* 3:65–71.
- COOPER, S. D., AND L. A. BARMUTA. 1993. Field experiments in biomonitoring. Pages 399–441 in D. M. Rosenberg and V. H. Resh (editors). *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman and Hall, New York.
- COSBY, B. J., G. M. HORNBERGER, J. N. GALLOWAY, AND R. F. WRIGHT. 1985. Modeling the effects of acid deposition: assessment of a lumped parameter model of soil water and streamwater chemistry. *Water Resources Research* 21:51–63.
- CRESSER, M. S., R. SMART, M. F. BILLET, C. SOULSBY, C. NEAL, A. WADE, S. LANGAN, AND A. C. EDWARDS. 2000. Modelling water chemistry for a major Scottish river from catchment attributes. *Journal of Applied Ecology* 37:171–184.
- CREUTZBURG, B. R., AND C. P. HAWKINS. 2008. What do J-NABS papers tell us about the state of knowledge in freshwater benthic science? *Journal of the North American Benthological Society* 27:593–604.
- CROKE, B. F. W., AND J. P. NORTON. 2004. Regionalisation of rainfall-runoff models. Pages 1201–1207 in *Transactions of the 2nd Biennial Meeting of the International Environmental Modelling and Software Society*. International Environmental Modelling and Software Society (iEMSs), Manno, Switzerland.

- CUNNANE, C. 1988. Methods and merits of regional flood frequency analysis. *Journal of Hydrology* 100:269–290.
- DADE, W. B., AND P. F. FRIEND. 1998. Grain-size, sediment-transport regime, and channel slope in alluvial rivers. *Journal of Geology* 106:661–675.
- DAVIS, W. M. 1890. The rivers of northern New Jersey, with notes on the classification of rivers in general. *National Geographic Magazine* 2:82–110.
- DAVY-BOWKER, J., R. T. CLARKE, R. K. JOHNSON, J. KOKES, J. F. MURPHY, AND S. ZAHRAĐKOVÁ. 2006. A comparison of the European Water Framework Directive physical typology and RIVPACS-type models as alternative methods of establishing reference conditions for benthic macroinvertebrates. *Hydrobiologia* 566: 91–105.
- DAYTON, P. K., M. J. TEGNER, P. B. EDWARDS, AND K. L. RISER. 1998. Sliding baselines, ghosts, and reduced expectations in kelp forest communities. *Ecological Applications* 8:309–322.
- DIXIT, S. S., J. P. SMOL, D. F. CHARLES, R. M. HUGHES, S. G. PAULSEN, AND G. B. COLLINS. 1999. Assessing water quality changes in the lakes of the northeastern United States using sediment diatoms. *Canadian Journal of Fisheries and Aquatic Sciences* 56:131–152.
- DODDS, W. K., AND R. M. OAKES. 2004. A technique for establishing reference nutrient concentrations across watersheds affected by humans. *Limnology and Oceanography: Methods* 2:333–341.
- DODDS, W. K., AND E. B. WELCH. 2000. Establishing nutrient criteria in streams. *Journal of the North American Benthological Society* 19:186–196.
- DODKINS, I., B. RIPPEY, T. J. HARRINGTON, C. BRADLEY, B. N. CHATHAIN, M. KELLY-QUINN, M. MCGARRIGLE, S. HODGE, AND D. TRIGG. 2005. Developing an optimal river typology for biological elements within the Water Framework Directive. *Water Research* 39: 3479–3486.
- DOLÉDEC, S., AND B. STATZNER. 2010. Responses of freshwater biota to human disturbances: contribution of *J-NABS* to developments in ecological integrity assessments. *Journal of the North American Benthological Society* 29:286–311.
- DOWNES, P. W. 1995. River channel classification for channel management purposes. Pages 347–365 in A. Gurnell and G. E. Petts (editors). *Changing river channels*. John Wiley and Sons, Chichester, UK.
- EATON, J. G., AND R. M. SCHELLER. 1996. Effects of climate warming on fish thermal habitat in streams of the United States. *Limnology and Oceanography* 41:1109–1115.
- EDINGER, J. E., D. W. DUTTWELER, AND J. C. GEYER. 1968. The response of water temperatures to meteorological conditions. *Water Resources Research* 4:1137–1143.
- EGGLISHAW, H., AND N. MORGAN. 1965. A survey of the bottom fauna of streams in the Scottish Highlands. Part II. The relationship of the fauna to the chemical and geological conditions. *Hydrobiologia* 26:173–183.
- EHRMAN, J. M., T. A. CLAIR, AND A. BOUCHARD. 1996. Using neural networks to predict pH changes in acidified eastern Canadian lakes. *AI Applications* 10:1–8.
- ERLANDSSON, M., J. FOLSTER, A. WILANDER, AND K. BISHOP. 2008. A metamodel based on MAGIC to predict the pre-industrial acidity status of surface waters. *Aquatic Sciences* 70:238–247.
- EUROPEAN COMMISSION. 2000. Directive 2000/60/EC of the European Parliament and the Council of the 23rd October 2000 establishing a framework for community action in the field of water policy. *Official Journal of the European Communities* L327: 1–72.
- EVANS, C. D., D. M. COOPER, S. JUGGINS, A. JENKINS, AND D. NORRIS. 2006. A linked spatial and temporal model of the chemical and biological status of a large, acid-sensitive river network. *Science of the Total Environment* 365:167–185.
- FEIO, M. J., S. F. P. ALMEIDA, S. C. CRAVEIRO, AND A. J. CALADO. 2009. A comparison between biotic indices and predictive models in stream water quality assessment based on benthic diatom communities. *Ecological Indicators* 9:497–507.
- FEIO, M. J., T. B. REYNOLDS, AND M. A. GRAÇA. 2006. Effect of seasonal changes on predictive model assessments of streams water quality with macroinvertebrates. *International Review of Hydrobiology* 91:509–520.
- FERGUSON, R. I., J. R. CUDDEN, T. B. HOEY, AND S. P. RICE. 2006. River system discontinuities due to lateral inputs: generic styles and controls. *Earth Surface Processes and Landforms* 31:1149–1166.
- FERNANDEZ, W., R. M. VOGEL, AND A. SANKARASUBRAMANIAN. 2000. Regional calibration of a watershed model. *Hydrological Sciences Journal – Journal des Sciences Hydrologiques* 45: 689–707.
- FERRÉOL, M., A. DOHET, H. CAUCHIE, AND L. HOFFMANN. 2005. A top-down approach for the development of a stream typology based on abiotic variables. *Hydrobiologia* 551:193–208.
- FLORES, A. N. 2004. Multi-scale models for predicting stream substrate size and channel morphology in watersheds of the Western US. MSc Thesis, Colorado State University, Fort Collins, Colorado.
- FRAPPIER, B., AND R. T. ECKERT. 2007. A new index of habitat alteration and a comparison of approaches to predict stream habitat conditions. *Freshwater Biology* 52:2009–2020.
- FRISSELL, C. A., W. J. LISS, C. E. WARREN, AND M. D. HURLEY. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. *Environmental Management* 10:199–214.
- FRYIRS, K. 2003. Guiding principles for assessing geomorphic river condition: application of a framework in the Bega catchment, South Coast, New South Wales, Australia. *Catena* 53:17–52.
- 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.
- GARDNER, B., P. J. SULLIVAN, AND A. J. LEMBO. 2003. Predicting stream temperatures: geostatistical model comparison using alternative distance metrics. *Canadian Journal of Fisheries and Aquatic Sciences* 60:344–351.
- GAYRAUD, S., B. STATZNER, P. Bady, A. HAYBACH, F. SCHÖLL, P. USSEGLIO-POLATERA, AND M. BACCHI. 2003. Invertebrate traits for the biomonitoring of large European rivers: an initial assessment of alternative metrics. *Freshwater Biology* 48:2045–2064.
- GELL, P., J. TIBBY, J. FLUIN, P. LEAHY, M. REID, K. ADAMSON, S. BULPIN, A. MACGREGOR, P. WALLBRINK, G. HANCOCK, AND B. WALSH. 2005. Accessing limnological change and variability using fossil diatom assemblages, South-East Australia. *River Research and Applications* 21:257–269.
- GERRITSEN, J. 1995. Additive biological indices for resource management. *Journal of the North American Benthological Society* 14: 451–457.
- GERRITSEN, J., M. T. BARBOUR, AND K. KING. 2000. Apples, oranges, and ecoregions: on determining pattern in aquatic assemblages. *Journal of the North American Benthological Society* 19: 487–496.
- GHERINI, S. A., L. MOK, R. J. M. HUDSON, G. F. DAVIS, C. W. CHEN, AND R. A. GOLDSTEIN. 1985. The ILWAS Model: formulation and application. *Water Air and Soil Pollution* 26:425–459.
- GIBSON, J. R., M. T. BARBOUR, J. B. STRIBLING, J. GERRITSEN, AND J. R. KARR. 1996. Biological criteria: technical guidance for streams and rivers (revised edition). EPA 822-B-96-001. Office of Water, US Environmental Protection Agency, Washington, DC.

- GILBERT, G. K. 1877. Report on the geology of the Henry Mountains. US Geographical and Geological Survey, Washington, DC.
- GOLDSTEIN, R. A., S. A. GHERINI, C. W. CHEN, L. MOK, AND R. J. M. HUDSON. 1984. Integrated acidification study (ILWAS): a mechanistic ecosystem analysis. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 305:409–425.
- GOMI, T., A. C. JOHNSON, R. L. DEAL, P. E. HENNON, E. H. ORLIKOWSKA, AND M. S. WIPFLI. 2006. Factors affecting distribution of wood, detritus, and sediment in headwater streams draining managed young-growth red alder–conifer forests in southeast Alaska. *Canadian Journal of Forest Research – Revue Canadienne de Recherche Forestière* 36:725–737.
- GOODERHAM, J. P. R., L. A. BARMUTA, AND P. E. DAVIES. 2007. Upstream heterogeneous zones: small stream systems structured by a lack of competence? *Journal of the North American Benthological Society* 26:365–374.
- GOTZINGER, J., AND A. BARDOSSY. 2007. Comparison of four regionalisation methods for a distributed hydrological model. *Journal of Hydrology* 333:374–384.
- GREENSTEIN, B., H. CURRAN, AND J. PANDOLFI. 1998. Shifting ecological baselines and the demise of *Acropora cervicornis* in the western North Atlantic and Caribbean Province: a Pleistocene perspective. *Coral Reefs* 17:249–261.
- GRENOUILLET, G., S. BROSE, L. TUDESQUE, S. LEK, Y. BARAILLÉ, AND G. LOOT. 2008. Concordance among stream assemblages and spatial autocorrelation along fragmented gradients. *Diversity and Distributions* 14:592–603.
- GUILLEMETTE, N., A. ST-HILAIRE, T. OUARDA, N. BERGERON, E. ROBICHAUD, AND L. BILODEAU. 2009. Feasibility study of a geostatistical modelling of monthly maximum stream temperatures in a multivariate space. *Journal of Hydrology* 364:1–12.
- GUSTARD, A. 1984. The characterisation of flow regimes for assessing the impact of water resource management on river ecology. Pages 53–60 in A. Lillehammer and S. J. Saltveit (editors). *Regulated rivers*. Universitetsforlaget As, Oslo, Norway.
- HARDING, J. S., M. J. WINTERBOURN, AND W. F. McDIFFETT. 1997. Stream faunas and ecoregions in South Island, New Zealand: do they correspond? *Archiv für Hydrobiologie* 140:289–307.
- HARGETT, E., J. ZUMBERGE, C. P. HAWKINS, AND J. OLSON. 2007. Development of a RIVPACS-type predictive model for bioassessment of Wadeable streams in Wyoming. *Ecological Indicators* 7:807–826.
- HAWKINS, C. P. 2006. Quantifying biological integrity by taxonomic completeness: its utility in regional and global assessments. *Ecological Applications* 16:1277–1294.
- HAWKINS, C. P., Y. CAO, AND B. ROPER. 2010. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. *Freshwater Biology* (in press). doi:10.1111/j.1365-2427.2009.02357.x
- HAWKINS, C. P., J. N. HOGUE, L. M. DECKER, AND J. W. FEMINELLA. 1997. Channel morphology, water temperature, and assemblage structure of stream insects. *Journal of the North American Benthological Society* 16:728–749.
- HAWKINS, C. P., J. L. KERSHNER, P. A. BISSON, M. D. BRYANT, L. M. DECKER, S. V. GREGORY, D. A. McCULLOUGH, C. K. OVERTON, G. H. REEVES, R. J. STEEDMAN, AND M. K. YOUNG. 1993. A hierarchical approach to classifying stream habitat features. *Fisheries* 18(6):3–12.
- HAWKINS, C. P., R. H. NORRIS, J. GERRITSEN, R. M. HUGHES, S. K. JACKSON, R. K. JOHNSON, AND R. J. STEVENSON. 2000a. Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. *Journal of the North American Benthological Society* 19:541–556.
- HAWKINS, C. P., R. H. NORRIS, J. N. HOGUE, AND J. W. FEMINELLA. 2000b. Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications* 10:1456–1477.
- HEERDEGEN, R., AND B. REICH. 1974. Unit hydrographs for catchments of different sizes and dissimilar regions. *Journal of Hydrology* 22:143–153.
- HEINO, J., AND H. MYKRÄ. 2006. Assessing physical surrogates for biodiversity: do tributary and stream type classifications reflect macroinvertebrate assemblage diversity in running waters? *Biological Conservation* 129:418–426.
- HENRIKSEN, A., J. KAMARI, M. POSCH, AND A. WILANDER. 1992. Critical loads of acidity: Nordic surface waters. *Ambio* 21:356–363.
- HENRIKSEN, J. A., J. HEASLEY, J. G. KENNEN, AND S. NIESWAND. 2006. Users' manual for the Hydroecological Integrity Assessment Process software (including the New Jersey Assessment Tools). US Geological Survey, Reston, Virginia.
- HERING, D., A. BUFFAGNI, O. MOOG, L. SANDIN, M. SOMMERHÄUSER, I. STUBAUER, C. FELD, R. JOHNSON, P. PINTO, N. SKOULIKIDIS, P. VERDONCHOT, AND S. ZAHRAĐKOVÁ. 2003. The development of a system to assess the ecological quality of streams based on macroinvertebrates – design of the sampling programme within the AQEM project. *International Review of Hydrobiology* 88:345–361.
- HERLIHY, A. T., R. M. HUGHES, AND J. C. SIFNEOS. 2006. National clusters of fish species assemblages in the conterminous United States and their relationship to existing landscape classification schemes. Pages 87–112 in R. M. Hughes, L. Wang, and P. W. Seelbach (editors). *Influences of landscapes on stream habitats and ecological assemblages*. American Fisheries Society, Bethesda, Maryland.
- HERLIHY, A. T., S. G. PAULSEN, J. VAN SICKLE, J. L. STODDARD, C. P. HAWKINS, AND L. L. YUAN. 2008. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. *Journal of the North American Benthological Society* 27:860–877.
- HERLIHY, A. T., AND J. C. SIFNEOS. 2008. Developing nutrient criteria and classification schemes for Wadeable streams in the conterminous US. *Journal of the North American Benthological Society* 27:932–948.
- HEUVELMANS, G., B. MUYS, AND J. FEYEN. 2006. Regionalisation of the parameters of a hydrological model: comparison of linear regression models with artificial neural nets. *Journal of Hydrology* 319:245–265.
- HIGGINS, J. V., M. T. BRYER, M. L. KHOURY, AND T. W. FITZHUGH. 2005. A freshwater classification approach for biodiversity conservation planning. *Conservation Biology* 19:432–445.
- HILL, T., P. WHITEHEAD, AND C. NEAL. 2002. Modelling long-term stream acidification in the chemically heterogeneous Upper Severn catchment, Mid-Wales. *Science of the Total Environment* 286:215–232.
- HOEY, T. B., AND B. J. BLUCK. 1999. Identifying the controls over downstream fining of river gravels. *Journal of Sedimentary Research* 69:40–50.
- HOLMES, M. G. R., A. R. YOUNG, A. GUSTARD, AND R. GREW. 2002. A region of influence approach to predicting flow duration curves within ungauged catchments. *Hydrology and Earth System Sciences* 6:721–731.
- HOOPER, R. P., N. CHRISTOPHERSEN, AND N. E. PETERS. 1990. Modeling streamwater chemistry as a mixture of soilwater end-members: an application to the Panola Mountain Catchment, Georgia, USA. *Journal of Hydrology* 116:321–343.
- HORNBERGER, G. M., B. J. COSBY, AND R. F. WRIGHT. 1989. Historical reconstructions and future forecasts of regional surface water

- acidification in southernmost Norway. *Water Resources Research* 25:2009–2018.
- HORTON, R. E. 1945. Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Geological Society of America Bulletin* 56: 275–370.
- HOSE, G., E. TURAK, AND N. WADDELL. 2004. Reproducibility of AUSRIVAS rapid bioassessments using macroinvertebrates. *Journal of the North American Benthological Society* 23: 126–139.
- HUGHES, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31–47 in W. S. Davis and T. P. Simon (editors). *Biological assessment and criteria*. Lewis Publishers, Boca Raton, Florida.
- HUGHES, R. M., S. A. HEISKARY, W. J. MATTHEWS, AND C. O. YODER. 1994. Use of ecoregions in biological monitoring. Pages 125–151 in S. L. Loeb and A. Spacie (editors). *Biological monitoring of aquatic systems*. Lewis Publishers, Boca Raton, Florida.
- HUGHES, R. M., D. P. LARSEN, AND J. M. OMERNIK. 1986. Regional reference sites: a method for assessing stream potentials. *Environmental Management* 10:629–635.
- HURLBERT, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187–211.
- HYNES, H. B. N. 1975. The stream and its valley. *Verhandlungen der Internationalen Vereinigung für theoretische und angewandte Limnologie* 19:1–15.
- IMMERZEEL, W. W., AND P. DROOGERS. 2008. Calibration of a distributed hydrological model based on satellite evapotranspiration. *Journal of Hydrology* 349:411–424.
- INFANTE, D. M., J. D. ALLAN, S. LINKE, AND R. H. NORRIS. 2009. Relationship of fish and macroinvertebrate assemblages to environmental factors: implications for community concordance. *Hydrobiologia* 623:87–103.
- INGLIS, L., I. K. G. BOOTHROYD, AND G. BRIERLEY. 2008. Effectiveness of the river environment classification in the Auckland Region. *New Zealand Geographer* 64:181–193.
- IVANOV, V. Y., E. R. VIVONI, R. L. BRAS, AND D. ENTEKHABI. 2004. Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: a fully-distributed physically-based approach. *Journal of Hydrology* 298:80–111.
- JARBOE, J., AND C. HAAN. 1974. Calibrating a water yield model for small ungaged watersheds. *Water Resources Research* 10: 256–262.
- JENERETTE, G. D., J. LEE, D. W. WALLER, AND R. E. CARLSON. 2002. Multivariate analysis of the ecoregion delineation for aquatic systems. *Environmental Management* 29:67–75.
- JOHNSON, L. B., AND G. E. HOST. 2010. Recent developments in landscape approaches for the study of aquatic ecosystems. *Journal of the North American Benthological Society* 29:41–66.
- JOHNSON, N. M., G. E. LIKENS, F. H. BORMANN, D. W. FISHER, AND R. S. PIERCE. 1969. Working model for the variation in stream water chemistry at the Hubbard Brook Experimental Forest, New Hampshire. *Water Resources Research* 5:1353–1363.
- JONES, K. L., G. C. POOLE, J. L. MEYER, W. BUMBACK, AND E. A. KRAMER. 2006. Quantifying expected ecological response to natural resource legislation: a case study of riparian buffers, aquatic habitat, and trout populations. *Ecology and Society* 11(2):15.
- JONES, M. L., C. K. MINNS, D. R. MARMOREK, AND F. C. ELDER. 1990. Assessing the potential extent of damage to inland lakes in eastern Canada due to acidic deposition: II. Application of the regional model. *Canadian Journal of Fisheries and Aquatic Sciences* 47:67–80.
- JOY, M. K., AND R. G. DEATH. 2002. Predictive modelling of freshwater fish as a biomonitoring tool in New Zealand. *Freshwater Biology* 47:2261–2275.
- JOY, M. K., AND R. G. DEATH. 2003. Assessing biological integrity using freshwater fish and decapod habitat selection functions. *Environmental Management* 32:747–759.
- JOY, M. K., AND R. G. DEATH. 2005. Modelling of fish and macrocrustacean assemblages for biological assessment in New Zealand. Pages 76–89 in S. Lek, M. Scardi, P. Verdonchot, J. P. Descy, and Y. S. Park (editors). *Modelling community structure in freshwater ecosystems*. Springer, Berlin, Germany.
- JUNCKER, P. 1971. Toward a global hydrological typology. *Journal of Hydrology (NZ)* 10:145–153.
- KARR, J. R. 1981. Assessment of biotic integrity using fish communities. *Fisheries* 6(6):21–27.
- KARR, J. R. 1991. Biological integrity: a long-neglected aspect of water resource management. *Ecological Applications* 1:66–84.
- KARR, J. R., AND E. W. CHU. 2000. Sustaining living rivers. *Hydrobiologia* 422:1–14.
- KAUFMANN, P. R., D. P. LARSEN, AND J. M. FAUSTINI. 2009. Bed stability and sedimentation associated with human disturbances in pacific northwest streams. *Journal of the American Water Resources Association* 45:434–459.
- KAY, A. L., D. A. JONES, S. M. CROOKS, A. CALVER, AND N. S. REYNARD. 2006. A comparison of three approaches to spatial generalization of rainfall-runoff models. *Hydrological Processes* 20: 3953–3973.
- KELLER, E. A., AND F. J. SWANSON. 1979. Effects of large organic material on channel form and fluvial processes. *Earth Surface Processes and Landforms* 4:361–380.
- KENNEN, J. G., L. J. KAUFFMAN, M. A. AYERS, D. M. WOLOCK, AND S. J. COLARULLO. 2008. Use of an integrated flow model to estimate ecologically relevant hydrologic characteristics at stream biomonitoring sites. *Ecological Modelling* 211:57–76.
- KILGOUR, B. W., K. M. SOMERS, AND D. E. MATTHEWS. 1998. Using the normal range as a criterion for ecological significance in environmental monitoring and assessment. *Ecoscience* 5: 542–550.
- KILGOUR, B. W., AND L. W. STANFIELD. 2006. Hindcasting reference conditions in streams. Pages 623–639 in R. M. Hughes, L. Wang, and P. W. Seelbach (editors). *Influences of landscapes on stream habitats and ecological assemblages*. American Fisheries Society, Bethesda, Maryland.
- KOKEŠ, J., S. ZAHŘÁDKOVÁ, D. NĚMEJCOVÁ, J. HODOVSKÝ, J. JARKOVSKÝ, AND T. SOLDÁN. 2006. The PERLA system in the Czech Republic: a multivariate approach for assessing the ecological status of running waters. *Hydrobiologia* 566:343–354.
- KONDOLF, G. M., D. R. MONTGOMERY, H. PIÉGAY, AND L. SCHMITT. 2003. Geomorphic classification of rivers and streams. Pages 169–202 in G. Kondolf and H. Piégay (editors). *Tools in fluvial geomorphology*. Wiley, New York.
- KRASOVSKAIA, I. 1997. Entropy-based grouping of river flow regimes. *Journal of Hydrology* 202:173–191.
- LANDRES, P. B., P. MORGAN, AND F. J. SWANSON. 1999. Overview of the use of natural variability concepts in managing ecological systems. *Ecological Applications* 9:1179–1188.
- LEATHWICK, J. R., D. ROWE, J. RICHARDSON, J. ELITH, AND T. HASTIE. 2005. Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology* 50:2034–2052.
- LENAT, D. R. 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. *Journal of the North American Benthological Society* 7: 222–233.
- LENAT, D. R. 1993. A biotic index for the southeastern United States: derivation and list of tolerance values, with criteria for assigning water-quality ratings. *Journal of the North American Benthological Society* 12:279–290.

- LEOPOLD, L. B., AND M. G. WOLMAN. 1957. River channel patterns: braided, meandering and straight. Professional Paper 282-B. US Geological Survey, Washington, DC.
- LEOPOLD, L. B., M. G. WOLMAN, AND J. P. MILLER. 1964. Fluvial processes in geomorphology. Dover Publications, New York.
- LEWIS, W. M., JR., J. M. MELACK, W. H. McDOWELL, M. MCCLAIN, AND J. E. RICHEY. 1999. Nitrogen yields from undisturbed watersheds in the Americas. *Biogeochemistry* 46:149–162.
- LINKE, S., R. C. BAILEY, AND J. SCHWINDT. 1999. Temporal variability of stream bioassessments using benthic macroinvertebrates. *Freshwater Biology* 42:575–584.
- LINKE, S., R. H. NORRIS, D. P. FAITH, AND D. STOCKWELL. 2005. ANNA: a new prediction method for bioassessment programs. *Freshwater Biology* 50:147–158.
- LORENZ, A., C. FELD, AND D. HERING. 2004. Typology of streams in Germany based on benthic invertebrates: ecoregions, zonation, geology and substrate. *Limnologica* 34:379–389.
- LOTSPEICH, F. B. 1980. Watersheds as the basic ecosystem: this conceptual framework provides a basis for a natural classification system. *Journal of the American Water Resources Association* 16:581–586.
- LYONS, J. 1996. Patterns in the species composition of fish assemblages among Wisconsin streams. *Environmental Biology of Fishes* 45:329–341.
- LYTLE, D. A., AND N. L. POFF. 2004. Adaptation to natural flow regimes. *Trends in Ecology and Evolution* 19:94–100.
- MARCHANT, R., F. WELLS, AND P. NEWALL. 2000. Assessment of an ecoregion approach for classifying macroinvertebrate assemblages from streams in Victoria, Australia. *Journal of the North American Benthological Society* 19:497–500.
- MASSONG, T. M., AND D. R. MONTGOMERY. 2000. Influence of sediment supply, lithology, and wood debris on the distribution of bedrock and alluvial channels. *Geological Society of America Bulletin* 112:591–599.
- MAXTED, J. R., M. T. BARBOUR, J. GERRITSEN, AND V. PORETTI. 2000. Assessment framework for mid-Atlantic coastal plain streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* 19:128–144.
- MAXWELL, J. R., C. J. EDWARDS, M. E. JENSEN, S. J. PAUSTIAN, H. PARROTT, AND D. M. HILL. 1995. A hierarchical framework of aquatic ecological units in North America (Nearctic Zone). General Technical Report NC-17. Forest Service, US Department of Agriculture, St. Paul, Minnesota.
- MAZOR, R. D., A. H. PURCELL, AND V. H. RESH. 2009. Long-term variability in bioassessments: a twenty-year study from two northern California streams. *Environmental Management* 43:1269–1286.
- MAZOR, R. D., T. B. REYNOLDSON, D. M. ROSENBERG, AND V. H. RESH. 2006. Effects of biotic assemblage, classification, and assessment method on bioassessment performance. *Canadian Journal of Fisheries and Aquatic Sciences* 63:394–411.
- MCBRIDE, G. B., J. C. LOFTIS, AND N. C. ADKINS. 1993. What do significance tests really tell us about the environment? *Environmental Management* 17:423–432.
- MCELRABY, E. P., G. A. LAMBERTI, AND V. H. RESH. 1989. Year-to-year variation in the aquatic macroinvertebrate fauna of a Northern California stream. *Journal of the North American Benthological Society* 8:51–63.
- MCINTYRE, N., H. LEE, H. WHEATER, A. YOUNG, AND T. WAGENER. 2005. Ensemble predictions of runoff in ungauged catchments. *Water Resources Research* 41:W12434.
- MEISNER, J. D. 1990. Effect of climatic warming on the southern margins of the native range of brook trout, *Salvelinus fontinalis*. *Canadian Journal of Fisheries and Aquatic Sciences* 47:1065–1070.
- MERRITT, R. W., D. H. ROSS, AND G. J. LARSON. 1982. Influence of stream temperature and seston on the growth and production of overwintering larval black flies (Diptera: Simuliidae). *Ecology* 63:1322–1331.
- MERZ, R., AND G. BLOSCHL. 2004. Regionalisation of catchment model parameters. *Journal of Hydrology* 287:95–123.
- MERZ, R., AND G. BLOSCHL. 2005. Flood frequency regionalization: spatial proximity vs. catchment attributes. *Journal of Hydrology* 302:283–306.
- MINSHALL, G. W. 1984. Aquatic insect–substratum relationships. Pages 358–400 in V. H. Resh and D. M. Rosenberg (editors). *The ecology of aquatic insects*. Praeger Publishers, New York.
- MINSHALL, G. W., AND J. N. MINSHALL. 1978. Further evidence on the role of chemical factors in determining the distribution of benthic invertebrates in the River Duddon. *Archiv für Hydrobiologie* 83:324–355.
- MIYAKE, Y., AND U. TAKEUCHI. 1951. On the temperature of river waters of Japan. *Japanese Journal of Limnology* 15:145–151.
- MOHSENI, O., H. G. STEFAN, AND J. G. EATON. 2003. Global warming and potential changes in fish habitat in US streams. *Climatic Change* 59:389–409.
- MOHSENI, O., H. G. STEFAN, AND T. R. ERICKSON. 1998. A nonlinear regression model for weekly stream temperatures. *Water Resources Research* 34:2685–2692.
- MONTGOMERY, D. R., AND J. M. BUFFINGTON. 1997. Channel-reach morphology in mountain drainage basins. *Geological Society of America Bulletin* 109:596–611.
- MONTGOMERY, D. R., W. E. DIETRICH, AND K. SULLIVAN. 1998. The role of GIS in watershed analysis. Pages 241–261 in S. N. Lane, K. S. Richards, and J. H. Chandler (editors). *Landform monitoring, modelling, and analysis*. John Wiley and Sons, Chichester, UK.
- MOOG, O., A. SCHMIDT-KLOIBER, T. OFENBÖCK, AND J. GERRITSEN. 2004. Does the ecoregion approach support the typological demands of the EU 'Water Framework Directive'? *Hydrobiologia* 516:21–33.
- MORIN, G., D. COUILLARD, D. CLUIS, H. G. JONES, AND J. M. GAUTHIER. 1987. Prediction of temperatures in rivers using a conceptual-model. *Hydrological Sciences Journal – Journal des Sciences Hydrologiques* 32:31–41.
- MOSLEY, M. 1981. Delimitation of New Zealand hydrologic regions. *Journal of Hydrology* 49:173–192.
- MOSS, D., M. T. FURSE, J. F. WRIGHT, AND P. D. ARMITAGE. 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biology* 17:41–52.
- MOSS, D., J. F. WRIGHT, M. T. FURSE, AND R. T. CLARKE. 2001. A comparison of alternative techniques for prediction of the fauna of running-water sites in Great Britain. *Freshwater Biology* 41:167–181.
- MUGODO, J., M. KENNARD, P. LISTON, S. NICHOLS, S. LINKE, R. H. NORRIS, AND M. LINTERMANS. 2006. Local stream habitat variables predicted from catchment scale characteristics are useful for predicting fish distribution. *Hydrobiologia* 572:59–70.
- MUNNÉ, A., AND N. PRAT. 2004. Defining river types in a Mediterranean area: a methodology for the implementation of the EU Water Framework Directive. *Environmental Management* 34:711–729.
- MUREȘAN, A. 2009. Relationship between the bed material size and the amount of metamorphic and volcanic rocks in hydrographic basins regarding two rivers from Maramureș Mountains (Eastern Carpathians – Romania). *Carpathian Journal of Earth and Environmental Sciences* 4:19–29.
- MYKRÄ, H., J. AROVITA, H. HÄMÄLÄINEN, J. KOTANEN, K.-M. VUORI, AND T. MUOTKA. 2008a. Assessing stream condition using macroinvertebrates and macrophytes: concordance of community

- responses to human impact. *Fundamental and Applied Limnology* 172/173:191–203.
- MYKRÄ, H., J. AROVITA, J. KOTANEN, H. HÄMÄLÄINEN, AND T. MUOTKA. 2008b. Predicting the stream macroinvertebrate fauna across regional scales: influence of geographical extent on model performance. *Journal of the North American Benthological Society* 27:705–716.
- MYKRÄ, H., J. HEINO, AND T. MUOTKA. 2007. Scale-related patterns in the spatial and environmental components of stream macroinvertebrate assemblage variation. *Global Ecology and Biogeography* 16:149–159.
- NAIMAN, R. J. 1998. Biotic stream classification. Pages 97–119 in R. J. Naiman and R. E. Bilby (editors). *River ecology and management: lessons from the Pacific Coastal ecoregion*. Springer-Verlag, New York.
- NANSON, G. C., AND J. C. CROKE. 1992. A genetic classification of floodplains. *Geomorphology* 4:459–486.
- NASH, J. E. 1960. Discussion, a unit hydrograph study, with particular reference to British catchments. *Proceedings of the Institution of Civil Engineers* 20:464–480.
- NAYEBI, M., D. KHALILI, S. AMIN, AND S. ZAND-PARSA. 2006. Daily stream flow prediction capability of artificial neural networks as influenced by minimum air temperature data. *Biosystems Engineering* 95:557–567.
- NELITZ, M. A., E. A. MACISAAC, AND R. M. PETERMAN. 2007. A science-based approach for identifying temperature-sensitive streams for rainbow trout. *North American Journal of Fisheries Management* 27:405–424.
- NERC (NATURAL ENVIRONMENT RESEARCH COUNCIL). 1975. Flood studies report. Institute of Hydrology, London, UK.
- NEWELL, R. L., AND G. W. MINSHALL. 1978. Life-history of a multivoltine mayfly, *Tricorythodes minutus*: 1 an example of effect of temperature on life cycle. *Annals of the Entomological Society of America* 71:876–881.
- NICHOLS, S. J., W. A. ROBINSON, AND R. H. NORRIS. 2006. Sample variability influences on the precision of predictive bioassessment. *Hydrobiologia* 572:215–233.
- NORRIS, R. H. 1995. Biological monitoring: the dilemma of data analysis. *Journal of the North American Benthological Society* 14:440–450.
- NORRIS, R. H., AND C. P. HAWKINS. 2000. Monitoring river health. *Hydrobiologia* 435:5–17.
- OBERDORFF, T., D. PONT, B. HUGUENY, AND D. CHESSEL. 2001. A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment. *Freshwater Biology* 46:399–415.
- OBERDORFF, T., D. PONT, B. HUGUENY, AND J. PORCHER. 2002. Development and validation of a fish-based index for the assessment of 'river health' in France. *Freshwater Biology* 47:1720–1734.
- ODE, P. R., C. P. HAWKINS, AND R. D. MAZOR. 2008. Comparability of biological assessments derived from predictive models and multimetric indices for increasing geographic scope. *Journal of the North American Benthological Society* 27:967–985.
- OLDEN, J. D. 2003. A species-specific approach to modeling biological communities and its potential for conservation. *Conservation Biology* 17:854–863.
- OLDEN, J. D., M. K. JOY, AND R. G. DEATH. 2006. Rediscovering the species in community-wide predictive modeling. *Ecological Applications* 16:1449–1460.
- OMERNIK, J. M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77:118–125.
- ORR, H. G., A. R. G. LARGE, M. D. NEWSON, AND C. L. WALSH. 2008. A predictive typology for characterising hydromorphology. *Geomorphology* 100:32–40.
- OSTERMILLER, J. D., AND C. P. HAWKINS. 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. *Journal of the North American Benthological Society* 23:363–382.
- UDIN, L., V. ANDREASSIAN, C. PERRIN, C. MICHEL, AND N. LE MOINE. 2008. Spatial proximity, physical similarity, regression and ungauged catchments: a comparison of regionalization approaches based on 913 French catchments. *Water Resources Research* 44:W03413.
- PAAVOLA, R., T. MUOTKA, R. VIRTANEN, J. HEINO, D. JACKSON, AND A. MAKI-PETÄYS. 2006. Spatial scale affects community concordance among fishes, benthic macroinvertebrates, and bryophytes in streams. *Ecological Applications* 16:368–379.
- PAAVOLA, R., T. MUOTKA, R. VIRTANEN, J. HEINO, AND P. KREIVI. 2003. Are biological classifications of headwater streams concordant across multiple taxonomic groups? *Freshwater Biology* 48:1912–1923.
- PAN, Y., R. J. STEVENSON, B. H. HILL, A. T. HERLIHY, AND G. B. COLLINS. 1996. Using diatoms as indicators of ecological conditions in lotic systems: a regional assessment. *Journal of the North American Benthological Society* 15:481–495.
- PAN, Y. D., R. J. STEVENSON, B. H. HILL, AND A. T. HERLIHY. 2000. Ecoregions and benthic diatom assemblages in Mid-Atlantic Highlands streams, USA. *Journal of the North American Benthological Society* 19:518–540.
- PANDEY, G. R., AND V. T. V. NGUYEN. 1999. A comparative study of regression based methods in regional flood frequency analysis. *Journal of Hydrology* 225:92–101.
- PARSONS, M., AND R. H. NORRIS. 1996. The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. *Freshwater Biology* 36:419–434.
- PAULY, D. 1995. Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology and Evolution* 10:430.
- PAUSTIAN, S. J., K. ANDERSON, D. BLANCHET, S. BRADY, M. CROPLEY, J. EDGINGTON, J. FRYXELL, G. JOHNEJACK, D. KELLIHER, M. KUEHN, S. MAKI, R. OLSON, J. SEESZ, AND M. WOLANEK. 1992. A channel type users guide for the Tongass National Forest, Southeast Alaska. Forest Service, US Department of Agriculture, Alaska Region. (Available from: http://www.fs.fed.us/r10/tongass/forest_facts/ct_guide/index.shtml)
- PEEL, M. C., F. H. S. CHIEW, A. W. WESTERN, AND T. A. McMAHON. 2000. Extension of unimpaired monthly streamflow data and regionalisation of parameter values to estimate streamflow in ungauged catchments. Centre for Environmental Applied Hydrology, Melbourne, Australia. (Available from: <http://www.anra.gov.au/topics/water/pubs/national/streamflow/streamflow.pdf>)
- PETERSON, E. E., A. A. MERTON, D. M. THEOBALD, AND N. S. URQUHART. 2006. Patterns of spatial autocorrelation in stream water chemistry. *Environmental Monitoring and Assessment* 121:571–596.
- PILGRIM, J. M., X. FANG, AND H. G. STEFAN. 1998. Stream temperature correlations with air temperatures in Minnesota: implications for climate warming. *Journal of the American Water Resources Association* 34:1109–1121.
- PINNENEGAR, J., AND G. ENGELHARD. 2008. The 'shifting baseline' phenomenon: a global perspective. *Reviews in Fish Biology and Fisheries* 18:1–16.
- POFF, N. L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society* 16:391–409.
- POFF, N. L., J. D. ALLAN, M. B. BAIN, J. R. KARR, K. L. PRESTEGAARD, B. D. RICHTER, R. E. SPARKS, AND J. C. STROMBERG. 1997. The natural flow regime. *BioScience* 47:769–784.
- POFF, N. L., J. D. OLDEN, N. K. M. VIEIRA, D. S. FINN, M. P. SIMMONS, AND B. C. KONDRATIEFF. 2006. Functional trait niches of North

- American lotic insects: traits-based ecological applications in light of phylogenetic relationships. *Journal of the North American Benthological Society* 25:730–755.
- PONT, D., B. HUGUENY, U. BEIER, D. GOFFAUX, A. MELCHER, R. NOBLE, C. ROGERS, N. ROSET, AND S. SCHMUTZ. 2006. Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages. *Journal of Applied Ecology* 43:70–80.
- POQUET, J. M., J. ALBA-TERCEDOR, T. PUNTÍ, M. DEL MAR SÁNCHEZ-MONTOYA, S. ROBLES, M. ÁLVAREZ, C. ZAMORA-MUÑOZ, C. E. SÁINZ-CANTERO, M. R. VIDAL-ABARCA, M. L. SUÁREZ, M. TORO, A. M. PUJANTE, M. RIERADEVALL, AND N. PRAT. 2009. The MEDiterranean Prediction and Classification System (MEDPACS): an implementation of the RIVPACS/AUSRIVAS predictive approach for assessing Mediterranean aquatic macroinvertebrate communities. *Hydrobiologia* 623:153–171.
- POWER, M. E., R. J. STOUT, C. E. CUSHING, P. P. HARPER, F. R. HAUER, W. J. MATTHEWS, P. B. MOYLE, B. STATZNER, AND I. R. WAIS DE BADGEN. 1988. Biotic and abiotic controls in river and stream communities. *Journal of the North American Benthological Society* 7: 456–479.
- PRINS, S. C. B., AND E. SMITH. 2007. Using biological metrics to score and evaluate sites: a nearest-neighbour reference condition approach. *Freshwater Biology* 52:98–111.
- RAHEL, F. J., AND N. P. NIBBELINK. 1999. Spatial patterns in relations among brown trout (*Salmo trutta*) distribution, summer air temperature, and stream size in Rocky Mountain streams. *Canadian Journal of Fisheries and Aquatic Sciences* 56:43–51.
- REECE, P. F., T. B. REYNOLDSON, J. S. RICHARDSON, AND D. M. ROSENBERG. 2001. Implications of seasonal variation for biomonitoring with predictive models in the Fraser River catchment, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences* 58:1411–1418.
- REHN, A. C., P. R. ODE, AND C. P. HAWKINS. 2007. Comparisons of targeted-riffle and reach-wide benthic macroinvertebrate samples: implications for date sharing stream-condition assessments. *Journal of the North American Benthological Society* 26: 332–348.
- REUSS, J. O., N. CHRISTOPHERSEN, AND H. M. SEIP. 1986. A critique of models for freshwater and soil acidification. *Water Air and Soil Pollution* 30:909–930.
- REYNOLDSON, T. B., R. C. BAILEY, AND R. H. NORRIS. 1995. Biological guidelines for freshwater sediment based on **B**enthic Assessment of **S**ediment (the BEAST) using multivariate approach for predicting biological state. *Australian Journal of Ecology* 20: 198–219.
- REYNOLDSON, T. B., R. H. NORRIS, V. H. RESH, K. E. DAY, AND D. M. ROSENBERG. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society* 16:833–852.
- RICE, S. 1998. Which tributaries disrupt downstream fining along gravel-bed rivers? *Geomorphology* 22:39–56.
- RICE, S. 1999. The nature and controls on downstream fining within sedimentary links. *Journal of Sedimentary Research* 69:32–39.
- RIEMAN, B. E., AND J. D. MCINTYRE. 1995. Occurrence of bull trout in naturally fragmented habitat patches of varied size. *Transactions of the American Fisheries Society* 124:285–296.
- ROBERTSON, D. M., D. A. SAAD, AND D. M. HEISEY. 2006. A regional classification scheme for estimating reference water quality in streams using land-use-adjusted spatial regression-tree analysis. *Environmental Management* 37:209–229.
- ROHM, C. M., J. M. OMERNIK, A. J. WOODS, AND J. L. STODDARD. 2002. Regional characteristics of nutrient concentrations in streams and their application to nutrient criteria development. *Journal of the American Water Resources Association* 38:213–239.
- ROSGEN, D. L. 1994. A classification of natural rivers. *Catena* 22: 169–199.
- ROSGEN, D. L. 1996. Applied river morphology. Wildland Hydrology, Pagosa Springs, Colorado.
- ROSS, G. A. 1970. The Stanford Watershed Model: the correlation of parameter values selected by a computerized procedure with measurable physical characteristics of the watershed. Water Resources Institute, University of Kentucky, Lexington, Kentucky.
- SANBORN, S. C., AND B. P. BLEDSOE. 2006. Predicting streamflow regime metrics for ungauged streams in Colorado, Washington, and Oregon. *Journal of Hydrology* 325:241–261.
- SÁNCHEZ-MONTOYA, M. D., T. PUNTÍ, M. L. SUÁREZ, M. D. VIDAL-ABARCA, M. RIERADEVALL, J. M. POQUET, C. ZAMORA-MUÑOZ, S. ROBLES, M. ÁLVAREZ, J. ALBA-TERCEDOR, M. TORO, A. M. PUJANTE, A. MUNN, AND N. PRAT. 2007. Concordance between ecotypes and macroinvertebrate assemblages in Mediterranean streams. *Freshwater Biology* 52:2240–2255.
- SANDIN, L., AND R. K. JOHNSON. 2000. Ecoregions and benthic macroinvertebrate assemblages of Swedish streams. *Journal of the North American Benthological Society* 19:462–474.
- SCARSBROOK, M. R. 2002. Persistence and stability of lotic invertebrate communities in New Zealand. *Freshwater Biology* 47:417–431.
- SCHECHER, W. D., AND C. T. DRISCOLL. 1988. An evaluation of the equilibrium calculations within acidification models: the effect of uncertainty in measured chemical components. *Water Resources Research* 24:533–540.
- SCHMITT, L., G. MAIRE, P. NOBELIS, AND J. HUMBERT. 2007. Quantitative morphodynamic typology of rivers: a methodological study based on the French Upper Rhine Basin. *Earth Surface Processes and Landforms* 32:1726–1746.
- SCHNOOR, J. L., AND W. STUMM. 1986. The role of chemical weathering in the neutralization of acidic deposition. *Schweizerische Zeitschrift für Hydrologie – Swiss Journal of Hydrology* 48:171–195.
- SCHUMM, S. A. 1977. The fluvial system. John Wiley and Sons, New York.
- SEELBACH, P. W., M. J. WILEY, M. E. BAKER, AND K. E. WEHRLY. 2006. Initial classification of river valley segments across Michigan's Lower Peninsula. *American Fisheries Society Symposium* 48: 25–48.
- SEFTON, C. E. M., AND S. M. HOWARTH. 1998. Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. *Journal of Hydrology* 211:1–16.
- SEFTON, C. E. M., P. G. WHITEHEAD, A. EATHERALL, I. G. LITTLEWOOD, AND A. J. JAKEMAN. 1995. Dynamic-response characteristics of the Plynlmon catchments and preliminary-analysis of relationships to physical descriptors. *EnvironMetrics* 6:465–472.
- SEIBERT, J. 1999. Regionalisation of parameters for a conceptual rainfall-runoff model. *Agricultural and Forest Meteorology* 98/ 99:279–293.
- SHAFFER, P., M. KENTULA, AND S. GWIN. 1999. Characterization of wetland hydrology using hydrogeomorphic classification. *Wetlands* 19:490–504.
- SHIELDS, A. 1936. Anwendung der Aehnlichkeitsmechanik und der Turbulenzforschung auf die Geschiebebewegung. *Mitteilungen der Preussischen Versuchsanstalt für Wasserbau und Schiffbau*. Heft 26. Preussischen Versuchsanstalt für Wasserbau und Schiffbau, Berlin, Germany.
- SIMPSON, G., E. SHILLAND, J. WINTERBOTTOM, AND J. KEAY. 2005. Defining reference conditions for acidified waters using a modern analogue approach. *Environmental Pollution* 137:119–133.
- SIMPSON, J. C., AND R. H. NORRIS. 2000. Biological assessment of river quality: development of AUSRIVAS models and outputs.

- Pages 125–142 in J. F. Wright, D. W. Sutcliffe, and M. T. Furse (editors). Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, UK.
- SINGH, K. P. 1971. Model flow duration and streamflow variability. *Water Resources Research* 7:1031–1036.
- SIVAPALAN, M., K. TAKEUCHI, S. W. FRANKS, V. K. GUPTA, H. KARAMBIRI, V. LAKSHMI, X. LIANG, J. J. McDONNELL, E. M. MENDIONDO, P. E. O'CONNELL, T. OKI, J. W. POMEROY, D. SCHERTZER, S. UHLENBROOK, AND E. ZEHE. 2003. IAHS decade on Predictions in Ungauged Basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal – Journal des Sciences Hydrologiques* 48:857–880.
- SKOIEN, J. O., R. MERZ, AND G. BLOCHL. 2006. Top-kriging: geostatistics on stream networks. *Hydrology and Earth System Sciences* 10:277–287.
- SLOANE, P. I. W., AND R. H. NORRIS. 2003. Relationship of AUSRIVAS-based macroinvertebrate predictive model outputs to a metal pollution gradient. *Journal of the North American Benthological Society* 22:457–471.
- SMALL, M. J., AND M. C. SUTTON. 1986. A direct distribution model for regional aquatic acidification. *Water Resources Research* 22:1749–1758.
- SMART, R. P., C. SOULSBY, M. S. CRESSER, A. J. WADE, J. TOWNEND, M. F. BILLETT, AND S. LANGAN. 2001. Riparian zone influence on stream water chemistry at different spatial scales: a GIS-based modelling approach, an example for the Dee, NE Scotland. *Science of the Total Environment* 280:173–193.
- SMART, R. P., C. SOULSBY, C. NEAL, A. WADE, M. S. CRESSER, M. F. BILLETT, S. J. LANGAN, A. C. EDWARDS, H. P. JARVIE, AND R. OWEN. 1998. Factors regulating the spatial and temporal distribution of solute concentrations in a major river system in NE Scotland. *Science of the Total Environment* 221:93–110.
- SMITH, J. G., J. J. BEAUCHAMP, AND A. J. STEWART. 2005. Alternative approach for establishing acceptable thresholds on macroinvertebrate community metrics. *Journal of the North American Benthological Society* 24:428–440.
- SMITH, K. 1981. The prediction of river water temperature. *Hydrological Sciences – Bulletin des Sciences Hydrologiques* 26:19–32.
- SMITH, M. J., W. R. KAY, D. H. D. EDWARD, P. J. PAPAS, K. S. J. RICHARDSON, J. C. SIMPSON, A. M. PINDER, D. J. CALE, P. H. J. HORWITZ, J. A. DAVIES, F. H. YUNG, R. H. NORRIS, AND S. A. HALSE. 1999. AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia. *Freshwater Biology* 41:269–282.
- SMITH, R. A., R. B. ALEXANDER, AND G. E. SCHWARZ. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous United States. *Environmental Science and Technology* 37:3039–3047.
- SNELDER, T. H., AND B. J. F. BIGGS. 2002. Multiscale River Environment Classification for water resources management. *Journal of the American Water Resources Association* 38:1225–1239.
- SNELDER, T. H., B. J. F. BIGGS, AND M. A. WEATHERHEAD. 2004a. Nutrient concentration criteria and characterization of patterns in trophic state for rivers in heterogeneous landscapes. *Journal of the American Water Resources Association* 40:1–13.
- SNELDER, T. H., B. J. F. BIGGS, AND R. A. WOODS. 2005. Improved eco-hydrological classification of rivers. *River Research and Applications* 21:609–628.
- SNELDER, T. H., F. CATTANÉO, A. M. SUREN, AND B. J. F. BIGGS. 2004b. Is the River Environment Classification an improved landscape-scale classification of rivers? *Journal of the North American Benthological Society* 23:580–598.
- SNELDER, T. H., K. DEY, AND J. LEATHWICK. 2007. A procedure for making optimal selection of input variables for multivariate environmental classifications. *Conservation Biology* 21:365–375.
- SNELDER, T. H., AND K. F. D. HUGHEY. 2005. The use of an ecologic classification to improve water resource planning in New Zealand. *Environmental Management* 36:741–756.
- SOBOTA, D. J., S. V. GREGORY, AND J. VAN SICKLE. 2006. Riparian tree fall directionality and modeling large wood recruitment to streams. *Canadian Journal of Forest Research – Revue Canadienne de Recherche Forestière* 36:1243–1254.
- SORANNO, P. A., K. S. CHERUVELIL, R. J. STEVENSON, S. L. ROLLINS, S. W. HOLDEN, S. HEATON, AND E. TORNG. 2008. A framework for developing ecosystem-specific nutrient criteria: integrating biological thresholds with predictive modeling. *Limnology and Oceanography* 53:773–787.
- STATZNER, B., P. Bady, S. DOLÉDEC, AND F. SCHÖLL. 2005. Invertebrate traits for the biomonitoring of large European rivers: an initial assessment of trait patterns in least impacted river reaches. *Freshwater Biology* 50:2136–2161.
- STEDINGER, J. R., AND G. D. TASKER. 1985. Regional hydrologic analysis 1. Ordinary, weighted, and generalized least-squares compared. *Water Resources Research* 21:1421–1432.
- STEFAN, H. G., AND E. B. PREUDHOMME. 1993. Stream temperature estimation from air-temperature. *Water Resources Bulletin* 29:27–45.
- STODDARD, J. L., D. P. LARSEN, C. P. HAWKINS, R. K. JOHNSON, AND R. H. NORRIS. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. *Ecological Applications* 16:1267–1276.
- STRAHLER, A. N. 1957. Quantitative analysis of watershed geomorphology. *American Geophysical Union* 38:913–920.
- STRIBLING, J. B., B. K. JESSUP, AND D. L. FELDMAN. 2008. Precision of benthic macroinvertebrate indicators of stream condition in Montana. *Journal of the North American Benthological Society* 27:58–67.
- STUMM, W., J. MORGAN, AND J. SCHNOOR. 1983. Acid rain, a consequence of man's alteration of hydrogeochemical cycles. *Naturwissenschaften* 70:216–223.
- SUTCLIFFE, D. W., AND T. R. CARRICK. 1973. Studies on mountain streams in the English Lake District. *Freshwater Biology* 3:437–462.
- SWANSON, F. J., G. W. LIENKAEMPER, AND J. R. SEDELL. 1976. History, physical effects, and management implications of large organic debris in western Oregon streams. General Technical Report PNW-GTR-56. Forest Service, US Department of Agriculture, Portland, Oregon.
- SWEENEY, B. W., AND R. L. VANNOTE. 1981. *Ephemerella* mayflies of White Clay Creek: bioenergetic and ecological relationships among six coexisting species. *Ecology* 62:1353–1369.
- TASKER, G. D. 1982. Comparing methods of hydrologic regionalization. *Water Resources Bulletin* 18:965–970.
- THEURER, F. D., K. A. VOOS, AND W. J. MILLER. 1984. Instream water temperature model. Instream Flow Information Paper 16. US Fish and Wildlife Service, US Department of Interior, Fort Collins, Colorado.
- THOMAS, D. M., AND M. A. BENSON. 1970. Generalization of stream-flow characteristics from drainage-basin characteristics. US Geological Survey, Washington, DC.
- THOMS, M. C., R. W. OGDEN, AND M. A. REID. 1999. Establishing the condition of lowland floodplain rivers: a palaeo-ecological approach. *Freshwater Biology* 41:407–423.
- TURAK, E., AND K. KOOP. 2008. Multi-attribute ecological river typology for assessing ecological condition and conservation planning. *Hydrobiologia* 603:83–104.

- UNDERWOOD, A. J. 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. *Ecological Applications* 4:3–15.
- USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2002. Summary of biological assessment programs and biocriteria development for states, tribes, territories, and interstate commissions: streams and Wadeable Rivers. EPA-822-R-02-048. US Environmental Protection Agency, Washington, DC.
- USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2006. Wadeable streams assessment: a collaborative survey of the Nation's streams. EPA 841-B-06-002. US Environmental Protection Agency, Washington, DC.
- VANNOTE, R. L., AND B. W. SWEENEY. 1980. Geographic analysis of thermal equilibria: a conceptual model for evaluating the effect of natural and modified thermal regimes on aquatic insect communities. *American Naturalist* 115:667–695.
- VAN SICKLE, J. 2008. An index of compositional dissimilarity between observed and expected assemblages. *Journal of the North American Benthological Society* 27:227–235.
- VAN SICKLE, J., AND S. V. GREGORY. 1990. Modeling inputs of large woody debris to streams from falling trees. *Canadian Journal of Forest Research – Revue Canadienne de Recherche Forestière* 20:1593–1601.
- VAN SICKLE, J., C. P. HAWKINS, D. P. LARSEN, AND A. T. HERLIHY. 2005. A null model for the expected macroinvertebrate assemblages in streams. *Journal of the North American Benthological Society* 24:178–191.
- VAN SICKLE, J., D. D. HUFF, AND C. P. HAWKINS. 2006. Selecting discriminant function models for predicting the expected richness of aquatic macroinvertebrates. *Freshwater Biology* 51:359–372.
- VAN SICKLE, J., AND R. M. HUGHES. 2000. Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon. *Journal of the North American Benthological Society* 19:370–384.
- VAN SICKLE, J., D. P. LARSEN, AND C. P. HAWKINS. 2007. Exclusion of rare taxa affects performance of the O/E index in bioassessments. *Journal of the North American Benthological Society* 26:319–331.
- VERDONSCHOT, P. F. M., AND R. C. NIJBOER. 2004. Testing the European stream typology of the Water Framework Directive for macroinvertebrates. *Hydrobiologia* 516:35–54.
- VIRTANEN, R., J. ILMONEN, L. PAASIVIRTA, AND T. MUOTKA. 2009. Community concordance between bryophyte and insect assemblages in boreal springs: a broad-scale study in isolated habitats. *Freshwater Biology* 54:1651–1662.
- WADE, A. J., B. M. JACKSON, AND D. BUTTERFIELD. 2008. Over-parameterised, uncertain 'mathematical marionettes' — How can we best use catchment water quality models? An example of an 80-year catchment-scale nutrient balance. *Science of the Total Environment* 400:52–74.
- WAGENER, T., AND H. S. WHEATER. 2006. Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. *Journal of Hydrology* 320:132–154.
- WAITE, I. R., A. T. HERLIHY, D. P. LARSEN, AND D. J. KLEMM. 2000. Comparing strengths of geographic and nongeographic classifications of stream benthic macroinvertebrates in the Mid-Atlantic Highlands, USA. *Journal of the North American Benthological Society* 19:429–441.
- WALKER, J. H., AND J. D. LAWSON. 1977. Natural stream temperature variations in a catchment. *Water Research* 11:373–377.
- WARD, J. C. 1963. Annual variation of stream water temperature. *American Society of Civil Engineers, Journal of the Sanitary Engineering Division* 89:3710–3732.
- WARFVINGE, P., AND H. SVERDRUP. 1992. Calculating critical loads of acid deposition with PROFILE—a steady-state soil chemistry model. *Water, Air, and Soil Pollution* 63:119–143.
- WEBB, B. W., D. M. HANNAH, R. D. MOORE, L. E. BROWN, AND F. NOBILIS. 2008. Recent advances in stream and river temperature research. *Hydrological Processes* 22:902–918.
- WEHRLY, K. E., T. O. BRENDEN, AND L. WANG. 2009. A comparison of statistical approaches for predicting stream temperatures across heterogeneous landscapes. *Journal of the American Water Resources Association* 45:986–997.
- WEHRLY, K. E., M. J. WILEY, AND P. W. SEELBACH. 2003. Classifying regional variation in thermal regime based on stream fish community patterns. *Transactions of the American Fisheries Society* 132:18–38.
- WEHRLY, K. E., M. J. WILEY, AND P. W. SEELBACH. 2006. Influence of landscape features on summer water temperatures in Lower Michigan streams. Pages 113–128 in R. M. Hughes, L. Wang, and P. W. Seelbach (editors). *Landscape influences on stream habitat and biological assemblages*. American Fisheries Society, Bethesda, Maryland.
- WHITEHEAD, P. G., E. J. WILSON, AND D. BUTTERFIELD. 1998. A semi-distributed Integrated Nitrogen model for multiple source assessment in Catchments (INCA): Part I — model structure and process equations. *Science of the Total Environment* 210:547–558.
- WILLOUGHBY, L. G., AND R. G. MAPPIN. 1988. The distribution of *Ephemera ignita* (Ephemeroptera) in streams: the role of pH and food resources. *Freshwater Biology* 19:145–155.
- WOHL, E., AND K. JAEGER. 2009. A conceptual model for the longitudinal distribution of wood in mountain streams. *Earth Surface Processes and Landforms* 34:329–344.
- 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.
- WRIGHT, J. F. 2000. An introduction to RIVPACS. Pages 1–24 in J. F. Wright, D. W. Sutcliffe, and M. T. Furse (editors). *Assessing the biological quality of fresh waters: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, Cumbria, UK.
- WRIGHT, J. F., D. MOSS, P. D. ARMITAGE, AND M. T. FURSE. 1984. A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. *Freshwater Biology* 14:221–256.
- YUAN, L. L. 2006. Theoretical predictions of observed to expected ratios in RIVPACS-type predictive model assessments of stream biological condition. *Journal of the North American Benthological Society* 25:841–850.
- YUAN, L. L., C. P. HAWKINS, AND J. VAN SICKLE. 2008. Effects of regionalization decisions on a O/E index for the US national assessment. *Journal of the North American Benthological Society* 27:892–905.
- ZHAI, J., C. T. DRISCOLL, T. J. SULLIVAN, AND B. J. COSBY. 2008. Regional application of the PnET-BGC model to assess historical acidification of Adirondack lakes. *Water Resources Research* 44:W01421. doi:10.1029/2006WR005532.
- ZHANG, Z. X., T. WAGENER, P. REED, AND R. BHUSHAN. 2008. Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization. *Water Resources Research* 44:W00B04. doi:10.1029/2008WR006833.

Received: 16 July 2009

Accepted: 10 November 2009