Development and comparison of stream indices of biotic integrity using diatoms vs. non-diatom algae vs. a combination

A. Elizabeth Fetscher, Rosalina Stancheva, J. Patrick Kociolek, Robert G. Sheath, Eric D. Stein, Raphael D. Mazor, Peter R. Ode, et al.

Journal of Applied Phycology

ISSN 0921-8971
Volume 26
Number 1

J Appl Phycol (2014) 26:433-450
DOI 10.1007/s10811-013-0088-2

Springer

J Appl Phycol (2014) 26:433–450
DOI 10.1007/s10811-013-0088-2

Development and comparison of stream indices of biotic integrity using diatoms vs. non-diatom algae vs. a combination

A. Elizabeth Fetscher · Rosalina Stancheva · J. Patrick Kociolek · Robert G. Sheath · Eric D. Stein · Raphael D. Mazor · Peter R. Ode · Lilian B. Busse

Received: 29 May 2013 / Revised and accepted: 11 July 2013 / Published online: 11 August 2013

Abstract Stream algal indices of biotic integrity (IBIs) are generally based entirely or largely on diatoms, because non-diatom ("soft") algae can be difficult to quantify and taxonomically challenging, thus calling into question their practicality and cost-effectiveness for use as bioindicators. Little has been published rigorously evaluating the strengths of diatom vs. soft algae-based indices, or how they compare to indices combining these assemblages. Using a set of ranked evaluation criteria, we compare indices of biotic integrity (IBIs) (developed for southern California streams) that incorporate different combinations of algal assemblages. We split a large dataset into independent "calibration" and "validation" subsets, then used the calibration subset to screen candidate metrics with respect to degree of responsiveness to anthropogenic stress, metric score distributions, and signal-to-noise ratio. The highest-performing metrics were combined into a total of 25 IBIs comprising either single-assemblage metrics (based on either diatoms or soft algae, including cyanobacteria) or combinations of metrics representing the two assemblages (for "hybrid IBIs"). Performance of all IBIs was assessed based on: responsiveness to anthropogenic stress (in terms of surrounding land uses and a composite water-chemistry gradient) using the validation data, and evaluated based on signal-to-noise ratio, metric redundancy, and degree of indifference to natural gradients. Hybrid IBIs performed best overall based on our evaluation. Single-assemblage IBIs ranked lower than hybrids vis-à-vis the abovementioned performance attributes, but may be considered appropriate for routine monitoring applications. Trade-offs inherent in the use of the different algal assemblages, and types of IBI, should be taken into consideration when designing an algae-based stream bioassessment program.

Keywords Freshwater diatoms · Soft algae · Cyanobacteria · Stream monitoring · Bioassessment · Index of biotic integrity (IBI)

Electronic supplementary material The online version of this article (doi:10.1007/s10811-013-0088-2) contains supplementary material, which is available to authorized users.

A. E. Fetscher (✉) · E. D. Stein · R. D. Mazor
Southern California Coastal Water Research Project, 3535 Harbor Boulevard, Suite 110, Costa Mesa, CA 92626, USA
e-mail: bettyf@sccwrp.org

R. Stancheva · R. G. Sheath
Department of Biological Sciences, California State University San Marcos, 333 S. Twin Oaks Valley Road, San Marcos, CA 92096, USA

J. P. Kociolek
Museum of Natural History and Department of Ecology and Evolutionary Biology, University of Colorado, UCB 218, Boulder, CO 80309, USA

R. D. Mazor · P. R. Ode
California Department of Fish and Wildlife, 2005 Nimbus Road, Rancho Cordova, CA 95670, USA

L. B. Busse
San Diego Regional Water Quality Control Board, 9174 Sky Park Court, Suite 100, San Diego, CA 92123, USA

Introduction

Stream bioassessment programs utilizing algal bioindicators are faced with the decision as to which assemblage(s) to include: diatoms, non-diatom ("soft") algae, or cyanobacteria. Together, these assemblages have the potential to offer multi-faceted characterization of water body condition and the stressors that may be affecting that condition. However, while there may be advantages to including algae from multiple assemblages, such an approach results in additional cost, training needs, and taxonomic expertise relative to what a single assemblage would require. Taxonomic needs can be offset somewhat by opting for coarser than species-level taxonomic resolution; however, the implications of this decision are not equally understood for all algal assemblages.

The array of choices in algal bioindicators leaves open questions about (1) the relative strengths of alternative algal indices, (2) the level of effort required to use them, and (3) how candidate indices relate to the monitoring questions of interest. Indices based on diatoms have been widely applied in stream bioassessment for decades (reviewed by Stevenson et al. 2010), and while the merit of soft algae (including cyanobacteria) as bioindicators has been noted by many investigators (Fernandez-Piñas et al. 1991; John and Johnson 1991; Pipp and Rott 1996; Douterelo et al. 2004; Rusanov et al. 2012; Whitton 2012) and several soft algae-based metrics have been described (Hill et al. 2000; Griffith et al. 2002; Porter et al. 2008), indices comprised solely of soft algae (Gutowski et al. 2004; Schaumburg et al. 2004; Schneider and Lindstrøm 2009, 2011) are comparatively rare. The discrepancy in the frequency of usage of soft algae indices relative to those using diatoms is likely due in part to challenges associated with soft algae species-level identification and precise quantification of specimens (Perona et al. 1998; Kelly et al. 2008; reviewed by Stancheva et al. 2012a), which have contributed to the impression that soft algae are less tractable and not cost-effective for use in bioassessment (Lavoie et al. 2004).

With some exceptions (e.g., Lavoie et al. 2004; Kelly 2006; Kelly et al. 2008; Schneider et al. 2012, 2013), surprisingly little has been published addressing the performance and/or relative strengths of diatom vs. soft algae as bioindicators, or how single-assemblage indices compare to those combining algal assemblages. Furthermore, while investigators have explored the value of generating diatom data based on varying levels of laboratory effort (e.g., comparing genus- vs. species-level identification; Hill et al. 2001; Lavoie et al. 2009) for bioassessment purposes, analogous studies are lacking for soft algae.

Here we report on development of a set of algae-based IBIs for use in southern California streams — classified in terms of (1) whether they are composed of a single assemblage (diatoms, or soft algae that include cyanobacteria) or a combination of diatoms and soft algae ("hybrids"), and (2) in the case of soft algae, the levels of effort (with respect to taxonomic resolution and specimen quantification) necessary for generating the data needed to calculate the different IBIs. Our goal is to compare performance among the different IBI types with respect to a series of ranked evaluation criteria in a way that can inform management decisions regarding development of stream algal bioassessment programs.

Methods

General approach to IBI development and comparison among IBIs

This study entailed the following: data collection and laboratory analyses; classifying sampling sites into "site disturbance classes"; splitting of the data into separate "calibration" and "validation" subsets; screening and scaling (and in a minority of cases, developing new; see below) metrics using the calibration subset of data; combining top-performing metrics into a set of IBIs consisting of diatoms only, soft algae only, or metrics from both assemblages (and requiring varying levels of laboratory effort); screening the IBIs for responsiveness to stress using the validation subset of data, and comparing the IBIs' relative performance with respect to a series of ranked evaluation criteria. The IBI development and evaluation process is summarized in a flow diagram in Online Resource 1.

A guiding principle in our IBI development was to base component metrics on existing knowledge about algal ecological traits. This goal was fully achievable with respect to diatoms, for which a wealth of published information on "ready-made" metrics is available for fine-tuning to local conditions. For a subset of soft algae metrics, however, inference about species' relationships to stream condition, based on the study's calibration dataset, was also employed. Nonetheless, even in this latter situation, literature was first consulted to identify the select set of parameters upon which to base the metrics. As such, metrics were based, at least to some degree, on a priori knowledge about ecological properties of algal taxa. This approach is in contrast to a purely statistical approach to bioassessment tool development in which the project dataset would be used to identify species relationships with the environment, upon which to create metrics de novo (e.g., Lavoie et al. 2006).

Data collection

The data used in IBI development came from the combined efforts of multiple monitoring programs using standardized field and laboratory procedures. The sampling sites utilized had been subject to a broad range of anthropogenic disturbances, but minimally disturbed "Reference" sites were also well represented in the combined data set, because several programs made an effort to collect samples at such locations. All told, 451 distinct stream reaches in California were sampled between 2007 and 2010. Some sites were sampled multiple times within a single site visit; data from replicates were used to evaluate metric and IBI signal-to-noise ratios.

Sampling sites were wadeable stream reaches delineated to be 150 m long (or in the less-common case of streams with a wetted width >10 m, reaches were 250 m long). Quantitative algae samples and various stream "physical habitat" measures (see below) were collected along a series of transects placed at equal intervals along the length of the reach. A "multihabitat method" was employed to objectively collect subsamples of algal specimens quantitatively from a known surface area over a representative sample of stream substrata. The algae subsamples were then composited, and aliquots were drawn from the composite for laboratory analysis. Across the length of the

delineated sampling reach, a non-quantitative (or "qualitative") soft algae sample was also generally collected. This consisted of collecting specimens of all macroalgal types that were observed in the stream reach during transit from one end to the other, yielding a species "inventory" of macroalgae for the reach. The macroalgal specimens were collected by hand, placed in Whirl-Pak bags containing native stream water, and stored on wet ice in the dark. Further details on all field procedures can be found in the report of Fetscher et al. (2009).

To assess the robustness of IBIs in the face of potential substratum effects on algal community composition (Rusanov et al. 2012), at a subset of sites (N=6), we also used a "targeted substratum" approach (Moulton et al. 2002): a quantitative sample was taken from each of the three most dominant substratum types, yielding four samples total (three different targeted samples plus the multihabitat sample described above). Additional data collected at each sampling site (one-time-only, concomitant with algae sampling) included water-chemistry constituents (nutrients, conductivity, pH, anions, dissolved organic carbon [DOC], dissolved metals) and physical habitat variables (canopy cover, gradient, pebble size distribution, riparian disturbance indicators) according to Fetscher et al. (2009).

Laboratory analyses

Diatom samples were cleaned according to the method of Van Der Werff (1955). For each sample, 600 valves were identified to below species level using oil immersion objectives (numerical aperture 1.40) at 1,000× magnification.

For soft algae, we employed a newly introduced method (Stancheva et al. 2012a) for laboratory processing in order to realize the full potential of this assemblage as bioindicators and allow us to assess the value of both high-resolution soft algae metrics and lower-resolution, depending upon whether the data utilized in metrics were quantitative or qualitative (i.e., species presence/absence), as well as whether species-level, or coarser, taxonomic resolution data were used in metric calculations.

Rather than homogenizing the entire original soft algae sample and using counting chambers, as is typical (Lowe and Laliberte 1996; Stevenson and Bahls 1999), in each sample, "macroalgae" (sensu Sheath and Cole 1992) were processed separately from microalgae. Macroalgae were removed from the original sample container gently with forceps, squeezed to remove as much liquid as possible, and then placed into a graduated centrifuge tube with a known volume of distilled water. Macroalgal total volume was determined by displacement, as indicated by the increase in volume (milliliters) of distilled water. The biovolume of each macroalgal species was then estimated under a stereomicroscope as its proportion of the total volume of macroalgae. In addition to collecting the biovolume information for each recorded macroalgal specimen, up to 100 non-diatom epiphytes were enumerated, if present.

A well-mixed 0.05-mL subsample from the remaining, microalgae-containing liquid in the sample was transferred to a standard microscope slide for viewing at 400×. At least 300 "natural algal counting entities" were identified and enumerated, and individual microscopical measurements were collected for each species along a known number of optical transects across the slide. The natural counting entity was defined as each naturally occurring form of algae (i.e., each unicell, filament, tissue-like form, coenocyte, colony, tuft, or crust) regardless of the number of cells in the thallus. The main purpose of using the concept of "counting entity" is to prevent numerous small cells in a sample with macroscopic forms from dominating a count relative to their actual contribution to the community biomass. It also facilitates the counting of algal forms that have linked cells that may be hard to distinguish. The separate processing of macroalgae and microalgae inherent in our procedure allowed identification of specimens to the lowest possible taxonomic level due to (1) the high-quality preservation of macroalgal vegetative and reproductive structures achievable because the samples were not homogenized prior to analysis, and (2) the even distribution and clear visualization of microalgal cells afforded by use of standard microscope slides.

The biovolume of each algal taxon encountered was calculated as individual biovolume (μm^3) per 1 cm^2 of stream-bottom area sampled. The resulting absolute biovolume of each algal taxon was then calculated as relative biovolume (in terms of the percentage of total algal biovolume represented by that individual taxon), for use in the biovolume-based soft algae metrics to be assessed for inclusion in the IBI(s).

In addition to the quantitative algal samples described above, "qualitative" samples of fresh, unfixed macroalgae were generally also collected from the sampling sites (see above). In the laboratory, material from the qualitative samples was scanned under dissecting and compound microscopes to identify each non-diatom macroalgal taxon to the lowest possible taxonomic level, resulting in a list of all macroalgal taxa in the qualitative sample.

Specimen observation and photomicrography were performed with an Olympus BX41 microscope and an Olympus SZ-40 stereo microscope with an attached Olympus MicroFire S99809 digital camera (Olympus Imaging America, USA). Further details on all soft algae laboratory procedures can be found in the work of Stancheva et al. (2012a, b).

Classifying sampling sites into disturbance classes

For index development, data from a large number of "reference" sites — those relatively unaffected by anthropogenic activities — are needed in addition to sites along a disturbance gradient representing a variety of stressors; reference sites

serve to set expectations for what biotic communities should look like with minimal human disturbance (Stoddard et al. 2006). We defined reference sites based largely on surrounding land use, but some local habitat data were included. In addition, upper limits for certain water-chemistry parameters (for total nitrogen and total phosphorus) were set. These values were used as "red flags" or "cross-checks" to alert us to potential anthropogenic stressors, not apparent from available land-use data, that could nonetheless be at play. They were set higher than what might be considered "typical" for Reference conditions in order to accommodate sites which, although essentially free from human influence, may experience relatively high nutrient concentrations due to non-anthropogenic factors, such as basin geology.

Landscape data were prepared by delineating the contributing watershed for each site from 30-m digital elevation models using a geographic information system, and clipping them at 5 and 1 km upstream of each site to facilitate assessment of disturbance at varying spatial scales. Metrics were then calculated from source layers relating to land cover, transportation structures, hydrology, and mining, and were used to assign sites to "disturbance classes" using the thresholds in Table 1.

All "Reference" thresholds had to be met in order for a site to be considered "Reference." We made a distinction among non-reference sites by classifying them as most "Disturbed" or "Intermediate," based on the same variables used to designate "Reference" status, but with more relaxed thresholds. All "Intermediate" thresholds had to be met in order for a site to be considered "Intermediate," and sites that met neither "Reference" nor "Intermediate" criteria automatically fell into "Disturbed."

Splitting the dataset into "calibration" and "validation" subsets

The project dataset was divided into subsets: 70 % of the sites (selected at random) were used for metric calibration (i.e., screening metrics with respect to their performance attributes and scaling them) and the remaining 30 % were set aside for IBI validation and comparison of relative performance of IBIs in terms of stressor–response.

Screening and scaling metrics

Metrics, grouped into "themes" organized by broader "categories" (Fore and Grafe 2002; Table 2), were screened for potential inclusion in IBIs. Most metrics were based on: (1) relative abundance of taxa that are indicators of the stream chemical environment or (2) relative abundance of taxa with morphological/behavioral characteristics rendering them differentially adapted to aspects of the stream physical environment (e.g., sedimentation tolerant, as indicated by taxa

Table 1 Thresholds for site-disturbance-class designations of stream sampling sites at various spatial scales

Variable	Scale	Threshold	
		Reference	Intermediate
Riparian disturbance (W1_Hall; Kaufmann et al. 1999)	local	1.5	3
% Agriculture	1 km, 5 km	3	30
	watershed	10	30
% Urban	1 km, 5 km	3	50
	watershed	10	30
% Agriculture + Urban	1 km, 5 km	5	–
% "Code 21"[a] land use	1 km, 5 km	7	50
	watershed	10	–
Road density (km km^{-2})	1 km, 5 km	2	10
	watershed	2	–
Road crossings (crossings/km^2; paved only)	1 km	5	–
	5 km	10	–
	watershed	50	–
Dam distance (km)	N/A	1	–
% Canals, pipes	watershed	10	50
Instream gravel mine density (mines/km stream)	5 km	0.1	–
No. producer mines	5 km	1	–
Total N (mg L^{-1})[b]	local	3	–
Total P (mg L^{-1})[c]	local	0.5	–

[a] "Code 21" encompasses a wide range of land uses primarily characterized by heavily managed vegetation (e.g., low-density residential development, parks, golf courses, highway medians)

[b,c] These values were used as "red flags" or "cross-checks" to alert us to potential anthropogenic stressors, not apparent from available land-use data, that could nonetheless be at play. They were set higher than what might be considered "typical" for Reference conditions in order to accommodate sites which, although essentially free from human influence, may experience relatively high nutrient concentrations due to non-anthropogenic factors, such as basin geology

Table 2 Categories and themes within which metrics were developed

Tolerance/ sensitivity	• Association with specific water-quality constituents (nutrients, organic carbon, metals)
	• Tolerant to low dissolved oxygen
	• Tolerant to high-ionic-strength/saline waters
Autecological guild	• Nitrogen fixers
	• Saprobic/heterotrophic taxa
Morphological guild	• Sedimentation indicators
Relationship to reference	• Taxa associated with reference vs. non-reference sites (Wang and Stevenson 2005)
Taxonomic groups	• Chlorophyta, Rhodophyta, Zygnemataceae, heterocystous cyanobacteria
Community form	• Total biovolume (soft algae)

motility; Bahls 1993). Various sources were consulted to attribute taxa for use in raw metric calculations. For diatoms, these included autecological information compiled by Spaulding et al. (2010) and by Porter et al. (2008), which in turn derived from sources including van Dam et al. (1994) and Potapova and Charles (2007). For soft algae, sources included Palmer (1969), Sládeček (1973), VanLandingham (1982), and Rott et al. (1997, 1999). However, autecological values were available for relatively few of our soft algae taxa. Therefore, in the case of some of our soft algae metrics, stressor relationships with specific taxa were of necessity derived empirically from the project calibration dataset. For the latter, we used indicator species analysis (Dufrêne and Legendre 1997) to identify taxa significantly associated with water-chemistry constituents previously shown to correlate with algal community attributes (Palmer 1969; Power 1990; Cattaneo et al. 1997; Vis et al. 1998; Leland and Porter 2000; Guasch et al. 2002; Komárek et al. 2002; Sheath 2003; Douterelo et al. 2004; Porter et al. 2008; John 2011; Stancheva et al. 2012b). These analyses were carried out on species absolute biovolume data using PC-ORD v6 software (McCune and Grace 2002). More details on this analysis are provided in Online Resource 2.

Diatom metrics were expressed in terms of proportion of valves (e.g., proportion of total valves belonging to *Epithemia* and *Rhopalodia*). In the case of soft algae, metrics were expressed in two ways: proportion of total species number, and proportion of total biovolume. Biovolume-based metrics were derived from the sum of the micro- and macroalgal components of each quantitative sample. For metrics based on species number, we availed ourselves to data not only from the quantitative field specimens for which biovolume values were calculated, but also from the epiphytes and the species recorded in the "qualitative" samples. This approach helped to mitigate one of the challenges inherent in using soft algae as bioindicators, specifically the fact that macroalgal forms are often patchily distributed in streams (Sheath et al. 1986), and therefore likely to be missed during more objective forms of sampling.

Using the calibration dataset, all raw metrics were subjected to a preliminary screen ("Phase 1") consisting of evaluation of data distributions and visualization of scatterplots depicting metric values along a composite landscape stressor gradient (Hering et al. 2006). The stressor gradient was constructed on the first principal component axis derived from a model using a subset of the same landscape variables listed in Table 1 (i.e., watershed-level percent agriculture, urban, and "Code 21" land uses, and road density). "Code 21" encompasses a wide range of land uses primarily characterized by heavily managed vegetation (e.g., low-density residential development, parks, golf courses, highway medians). If a given metric failed to exhibit the expected response to stress (see Online Resource 3, column 5), and/or had a large proportion of zeros (Stoddard et al. 2008), it was eliminated. Also eliminated were species-number-based soft algae metrics showing sensitivity to whether a qualitative sample had been available for analysis (based on *t*-tests of raw metric scores, using as the grouping variable whether or not a qualitative sample had been collected). Metrics passing this initial phase were scaled into standardized, unitless forms according to Ode et al. (2005).

For "Phase-2" metric screening, using the calibration data set, Spearman rank correlation was used to assess relationships between scaled metrics and a stressor gradient constructed from the first principal component axis derived from conventional water chemistry parameters (chloride, DOC, conductivity, and sulfate) that tend to increase with anthropogenic impacts. Signal-to-noise ratio was determined for each metric by comparing variance of each metric among streams (i.e., "signal") with variance between replicate samples collected at the same site (i.e., "noise") (Kaufmann et al. 1999) using restricted maximum likelihood (REML). Data distributions of the scaled metrics were visualized with histograms. Metrics exhibiting a poor distribution (e.g., strongly bimodal) were eliminated. Based on these screens, a "long list" of successful metrics was generated by giving preference to those exhibiting the strongest relationships with stress (Fore and Grafe 2002), followed by highest signal-to-noise ratios (Stoddard et al. 2008), then acceptable distributions. The best-performing metrics within each metric theme (Table 2), up to a total of two within each of the diatom and soft algae assemblages, were retained for incorporation into IBIs.

Combining metrics into IBIs

IBIs were created by summing different sets of five to ten long-list diatom and/or soft algae metrics that overlapped as little as possible in terms of metric themes (Table 2). Multipliers were used to scale each IBI to a maximum possible score of 100. The IBIs were divided into categories based on their metric composition: diatom only vs. soft algae only vs. a combination of both metric types (i.e., hybrid IBIs). Among the hybrids, those requiring execution of the full soft algae laboratory protocol were distinguished from those requiring only a subset of that effort, i.e., those requiring species-number information but not biovolume, and those based on biovolumes but not requiring species-level identification.

Validating the IBIs and comparing their relative performance

From this point (the validation stage) forward in the IBI development process, we focused our efforts on the southern California portion of the dataset, in order to produce IBIs specifically tuned for that region. This is because southern California is where the greatest density of data (i.e., the largest number of sites) was concentrated, and also because it was the

region within which the broadest gradients of human disturbance were captured by the available data. As such, unless otherwise noted, all reported performance characteristics past the initial metric screening phase focus only on the southern California subset of sites, and the resulting indices are intended specifically for application in that region of the state.

We validated the IBIs and compared their relative performance by subjecting them to a series of ranked screening criteria. Except where noted, in the case of screens dealing with stress response, only sites in the "validation" subset of data were used. IBI relationship with stress was accorded priority (Fore 2003), with signal-to-noise deemed the second most important criterion. If values for these priority screens exhibited minimal difference among multiple IBIs, then a pair of additional, lower-priority screen types — redundancy among metrics (in the form of mean correlation among metrics) and indifference to natural gradients — were also taken into consideration.

Using the validation data set, IBIs were evaluated for relationships to stressors via two approaches. One involved assessing how well IBI scores separated sites belonging to the Disturbed, Intermediate, and Reference site disturbance classes (Table 1) using ANOVA with Tukey's tests for multiple comparisons, along with visualization of box plots to evaluate overlap between interquartile ranges (Barbour et al. 1996; Klemm et al. 2002). The second approach involved determining Spearman rank correlations between IBI scores and the same water-chemistry principal component axis that was used in Phase 2 of metric screening. This latter analysis was conducted both on the validation sites across the full site disturbance gradient, and on the validation plus calibration sites within the Intermediate site-disturbance class (to assess the IBIs' ability to resolve sites exposed to intermediate disturbance levels).

Signal-to-noise ratio was determined for each IBI, similar to the method described for metric screening. However, in addition to evaluating signal-to-noise among true replicate samples, it was also evaluated across all the same-day samples collected at sites where both the multihabitat and targeted-substrata samplings were carried out.

To evaluate the level of redundancy of information among metrics (Cao et al. 2007), the mean Spearman's ρ value for all the pairwise metric combinations was calculated for each IBI (Van Sickle 2010).

Ideally, an IBI should be relatively indifferent to sources of natural variation, such that variation in IBI scores among sites is most likely the result of anthropogenic, rather than non-anthropogenic, factors. We evaluated this parameter using Spearman rank correlation to assess relationships between the IBIs and a large suite of (primarily) natural gradients to which algae might be responsive. These included: percent fines, percent fines + sand, alkalinity, reach- and landscape-level slope, canopy cover, stream order, watershed area, elevation, latitude, longitude, and selected climate variables (mean annual precipitation and maximum air temperature associated with the site, based on records from 1971 to 2000; PRISM Climate Group 2013). This analysis was conducted within the Reference group of sites only, in order to reduce the likelihood that any responsiveness realized might have an anthropogenic component (Cao et al. 2007; Schneider 2011).

Upon identifying the top-performing IBI within each assemblage/effort category, in order to facilitate a final comparison among IBIs, linear regression was used to visualize relationships between IBI scores and the water-chemistry principal component described previously.

Defining IBI scoring categories

For the top-performing IBI, similarly to the approaches used by Ode et al. (2005) and Schneider (2011), we used information on the standard deviation of IBI scores among Reference sites in order to create a means of classifying new sites according to their IBI scores. This was accomplished by establishing a statistical boundary below which IBI scores could be considered to be distinct from that associated with reference conditions. Our boundary was designated as two standard deviations below the mean Reference site IBI score within the project dataset.

Results

Overall, we classified 27 % of sites as "Reference," 38 % as "Intermediate," and 35 % as "Disturbed." In all cases, sites were excluded from the "Reference" classification based on one or more land-use or local riparian disturbance (W1_Hall) screens (Table 1). Reference sites occurred in many parts of the state, but their spatial density varied by region (Online Resource 4).

Metric screening and scaling

Of the 87 metrics tested (Online Resource 3), 56 % were excluded based on Phase 1 screening of raw metrics, and another 20 % during the Phase 2 screening of scaled metrics, resulting in 21 "long-list" metrics. Most eliminations in Phase 1 were due to poor distribution of metric scores and/or lack of relationship to the landscape stressor gradient, while a small subset were eliminated due to metric sensitivity to whether or not a qualitative sample had been collected. Poor distribution of scores was more common among the soft algae metrics than the diatoms, and occurred both in species-number and biovolume-based metrics. Taxon designations from the indicator species analysis are provided in Online Resource 2. The

Table 3 Metric composition of IBIs developed for use in southern California

IBI	Proportion highly motile (d)	Proportion sediment tolerant (highly motile) (d)	Proportion low N indicators (d)	Proportion low P indicators (d)	Proportion N heterotrophs (d)	Proportion requiring >50 % DO saturation (d)	Proportion requiring nearly 100 % DO saturation (d)	Proportion halobiontic (d)	Proportion oligo- and beta-mesosaprobic (d)	Proportion poly- and eutrophic (d)	Proportion *A. minutissimum* (d)
H10		×		×	×		×	×			
H13		×		×	×	×		×			
H14		×		×	×	×		×			
H15		×		×	×	×		×			
H16	×			×	×	×		×			
H17		×		×	×	×		×			
H18		×		×	×	×		×			
H19		×	×		×	×		×			
H2	×				×	×		×			
H20		×	×		×	×		×			
H21		×		×	×	×		×			
H22		×		×	×	×		×			
H23		×		×	×	×		×			
H3	×		×	×	×	×		×			
H6		×	×	×	×			×			
H7		×	×	×	×			×			
H9		×	×	×	×		×	×			
D13		×	×	×	×			×			
D14	×		×	×	×			×			
D16		×	×	×	×	×		×	×		×
D17	×				×		×	×	×	×	×
D18		×		×	×	×		×			
S1											
S11											
S2											

IBI	Proportion Chlorophyta (s, b)	Proportion of green algae belonging to CRUS (s, b)	Proportion ZHR (s, m)	Proportion ZHR (s, b)	Proportion high DOC indicators (s, b)	Proportion "non-reference" indicators (s, b)	Proportion high Cu indicators (s, sp)	Proportion high DOC indicators (s, sp)	Proportion low TP indicators (s, sp)	Proportion "non-reference" indicators (s, sp)
H10		×	×				×			
H13		×		×		×				
H14				×	×					
H15			×					×		
H16				×				×		
H17			×							×
H18				×		×				
H19					×		×		×	
H2			×				×	×	×	×
H20							×	×	×	

Table 3 (continued)

IBI	Proportion Chlorophyta (s, b)	Proportion of green algae belonging to CRUS (s, b)	Proportion ZHR (s, m)	Proportion ZHR (s, b)	Proportion high DOC indicators (s, b)	Proportion "non-reference" indicators (s, b)	Proportion high Cu indicators (s, sp)	Proportion high DOC indicators (s, sp)	Proportion low TP indicators (s, sp)	Proportion "non-reference" indicators (s, sp)
H21	×			×						
H22		×		×						
H23		×	×					×		
H3		×	×							
H6			×					×		
H7		×	×				×	×		
H9		×	×				×	×		
D13										
D14										
D16										
D17										
D18										
S1		×	×		×		×		×	×
S11		×		×		×	×		×	
S2		×	×				×	×	×	×

Note: Some of the IBIs appear across two rows, as opposed to just one. It is necessary to read through both pages of the entire table in order to capture all the metrics that are present in any given IBI.

IBI names beginning with "H" are hybrids, "D" are diatom-only, and "S" are soft algae-only. Metric names followed by "d" are derived from the diatom assemblage, and those followed by "s" are from the soft algae. Of the soft algae, those followed by: "b" are based on biovolume, "sp" are based on species presence, and "m" are the average of the "b" and "sp" counterpart metric values. "CRUS" stands for *Cladophora glomerata* + *Rhizoclonium hieroglyphicum* + *Ulva flexuosa* + *Stigeoclonium* spp. "ZHR" stands for Zygnemataceae + heterocystous cyanobacteria + Rhodophyta. "Green algae" refers to taxa within Chlorophyta + Charophyta

Table 4 IBI performance results by effort category

Effort category	IBI	Correlation[a] with PC1 (validation)	Correlation[b] with PC1 (within Intermediate class)	R^2, ANOVA with Site Disturbance Class (validation)	Are all pairwise[c] differences significant?	Signal-to-noise (multihabitat replicates)	Signal-to-noise (targeted substrata)	Metrics mean pairwise correlation	No. natural gradients with significant correlation
Single assemblage	**D18**	−0.71	−0.47	0.32	Yes	18.6	10.5	0.60	6
	D13	−0.68	−0.44	0.34	Yes	16.5	14.4	0.60	3
	D14	−0.68	−0.42	0.33	Yes	15.1	13.3	0.58	3
	D16	−0.66	−0.42	0.36	Yes	18.6	17.9	0.60	5
	D17	−0.65	−0.38	0.31	Yes	16.2	11.2	0.56	3
	S2	−0.57	−0.29	0.52	Not R vs. I	18.2	3.2	0.61	3
	S1	−0.50	−0.27	0.52	Not R vs. I	31.3	3.2	0.61	3
	S11	−0.43	−0.22	0.48	Not R vs. I	24.0	4.3	0.55	4
Full effort	**H23**	−0.69	−0.51	0.51	Yes	34.7	7.3	0.43	1
	H19	−0.71	−0.48	0.55	Not R vs. I	29.3	15.7	0.45	1
	H2	−0.67	−0.45	0.54	Not R vs. I	21.2	6.3	0.47	0
	H7	−0.70	−0.47	0.54	Yes	24.6	11.5	0.44	2
	H9	−0.68	−0.46	0.51	Yes	21.7	11.7	0.42	2
	H13	−0.67	−0.44	0.50	Not R vs. I	51.6	7.5	0.41	0
	H10	−0.68	−0.45	0.48	Yes	16.5	9.9	0.39	2
	H18	−0.65	−0.44	0.47	Not R vs. I	40.7	10.0	0.44	0
	H6	−0.68	−0.51	0.47	Yes	26.8	11.8	0.50	4
	H14	−0.66	−0.49	0.46	Not R vs. I	33.5	8.6	0.46	1
	H15	−0.71	−0.53	0.46	Yes	29.3	8.6	0.48	3
	H17	−0.70	−0.50	0.45	Yes	28.4	8.8	0.48	1
	H3	−0.66	−0.47	0.45	Yes	24.6	10.5	0.46	1
	H16	−0.69	−0.52	0.44	Yes	22.3	7.9	0.46	1
	H22	−0.67	−0.46	0.44	Yes	24.0	8.4	0.43	0
Reduced effort (no species level for soft)	**H21**	−0.63	−0.44	0.41	Yes	24.6	12.9	0.42	1
Reduced effort (no biovolume)	**H20**	−0.72	−0.51	0.51	Yes	20.7	12.0	0.50	1

Only data from sites in southern California are included in the analyses. IBIs in boldface type are analyzed in more detail

[a] All values significant (α=0.05)

[b] All values significant (α=0.05) except for S11 (P=0.08)

[c] "Not R vs. I" indicates that the Reference and Intermediate site disturbance class IBI scores were not significantly different

Table 5 Summary of performance outcomes for five algal IBI effort categories

Effort category	Stress response: correlation with PC1 (validation)	Stress response: correlation with PC1 (within Intermediate class)	Stress response: variance in IBI scores explained by Site Disturbance Class (validation)	Signal-to-noise (multihabitat replicates)	Minimization of metric redundancy	Indifference to natural gradients	Laboratory effort (as multiplier, relative to diatoms alone)
Reduced effort hybrid (no biovolume)	**1**	**1**	**1**	4	3	**1**	2
Full effort hybrid	2	2	2	**1**	2	2	2.5
Reduced effort hybrid (no species level, for soft algae)	3	3	3	2	**1**	**1**	2
Single assemblage (soft)	4	5	**1**	3	4	3	1.5[a]
Single assemblage (diatom)	2	4	4	5	4	4	**1**

Values in the first six columns are ranks based on mean performance outcomes (from Table 4) across IBIs. Final column is the estimated relative laboratory effort, expressed as a multiple of the amount of labor diatoms alone would require. Data in boldface indicate the top-performing IBI effort category(ies) for each criterion. Notes: (1) for some criteria, differences in performance between IBI effort categories were relatively small; (2) only one IBI was available for each of the two reduced-effort hybrid IBI categories

[a] All IBIs based on soft algae alone required execution of the *full* soft algae laboratory protocol (Stancheva et al. 2012a)

values used for scaling metrics are provided in Online Resources 5 (diatoms) and 6 (soft algae).

IBI development and validation

Twenty-five IBIs were developed (Table 3). In general, the hybrid IBIs outperformed the single-assemblage IBIs based on responsiveness to stress (Tables 4 and 5). Hybrids were best able to discriminate between site-disturbance classes based on IBI score distributions, and interquartile ranges for scores of top-performing hybrids exhibited less overlap between site-disturbance classes than their single-assemblage counterparts. The top-performing soft algae IBI exhibited substantial separation between the Disturbed and Intermediate site classes but little separation between Intermediate and Reference, whereas the opposite was true for the top-performing diatom IBI (Fig. 1).

Signal-to-noise ratio for replicates collected via the multihabitat field protocol was consistently higher than that resulting from sampling different targeted substrata (Table 4). On average, hybrids outperformed single-assemblage IBIs (particularly the diatoms) for the replicate multihabitat sampling, and soft algae-only IBIs exhibited the lowest signal-to-noise ratio among the targeted-substrata samples. Mean pairwise correlation coefficients among metrics across all IBIs ranged from 0.39 to 0.61, and were invariably lower among hybrids than among single-assemblage IBIs (Tables 4 and 5). IBIs varied considerably in terms of their relationships to natural gradients, with some not significantly correlated with any of the 13 factors tested, and others correlated with ≥4 (Table 4). The natural gradient most commonly significantly correlated with IBI scores was watershed area (positively associated with 17 hybrid and diatom IBIs) followed by stream order and percent fines (correlated with seven IBIs each, positively for stream order and negatively for fines). Similarly, Schneider (2011) noted a significant effect of catchment size on scores of the acidification index periphyton (AIP) among reference streams in Norway, albeit the AIP is based entirely on soft algae. Overall, our hybrid IBIs were least frequently correlated with the natural gradients tested, whereas diatom-only IBIs exhibited significant relationships with the highest number of natural gradients (Table 4). Diatom-only and soft algae-only IBIs responded differently to natural gradients. Diatom IBIs were particularly responsive to stream order, watershed area, and percent fines, and soft algae IBIs were most responsive to canopy cover and slope (both negatively).

Comparison of top-performing IBIs across effort categories

Based on the information in Table 4, the top performing IBI from each of the effort categories was selected for further evaluation. These included three hybrids representing different

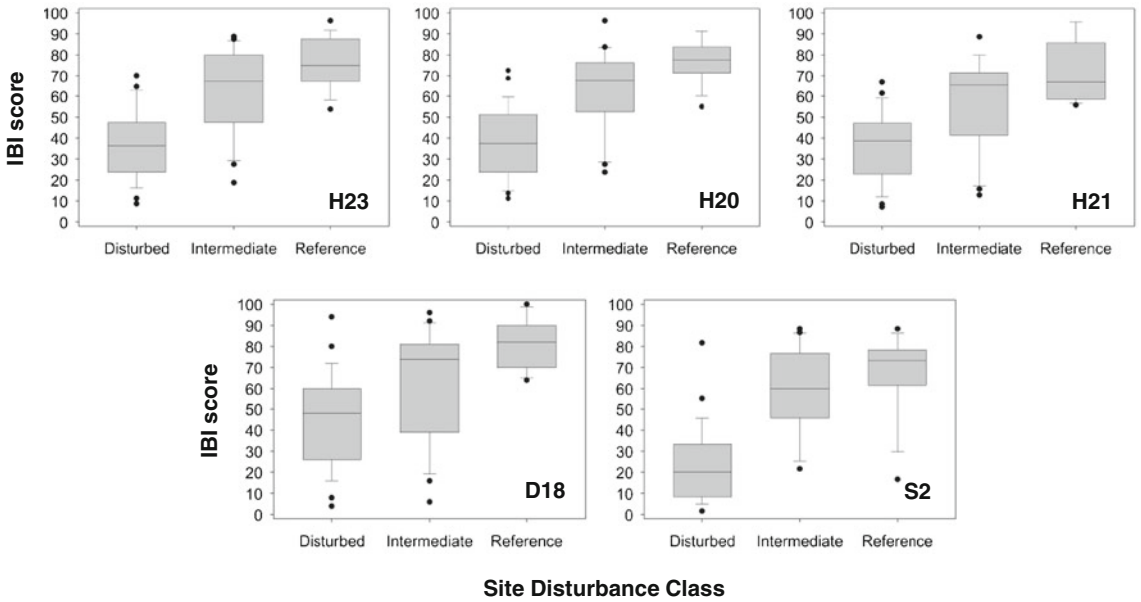

Fig. 1 Discriminatory power of IBIs from different effort categories. Data are from southern California sites within the validation dataset

levels and types of laboratory effort (H23: requiring the full soft algae laboratory protocol, H20: requiring no biovolume data, but species-level identification of soft algae taxa, and H21: requiring soft algae biovolumes, but genus-level or above identifications of soft algae taxa), and one single-assemblage IBI (D18). The performance characteristics of these four IBIs are highlighted in Figs. 2–5.

All four IBIs were responsive to stress. Slopes of the regressions of IBI scores on the water-chemistry principal component were similar among the four IBIs, ranging from −13.6 to −11.7. R^2 values for the regressions, in descending order, were 0.504 (for IBI H20), 0.485 (H23), 0.437 (D18), and 0.394 (H21) (Fig. 2). Within the Intermediate site-disturbance class, slopes ranged from −7.24 to −6.1, and R^2

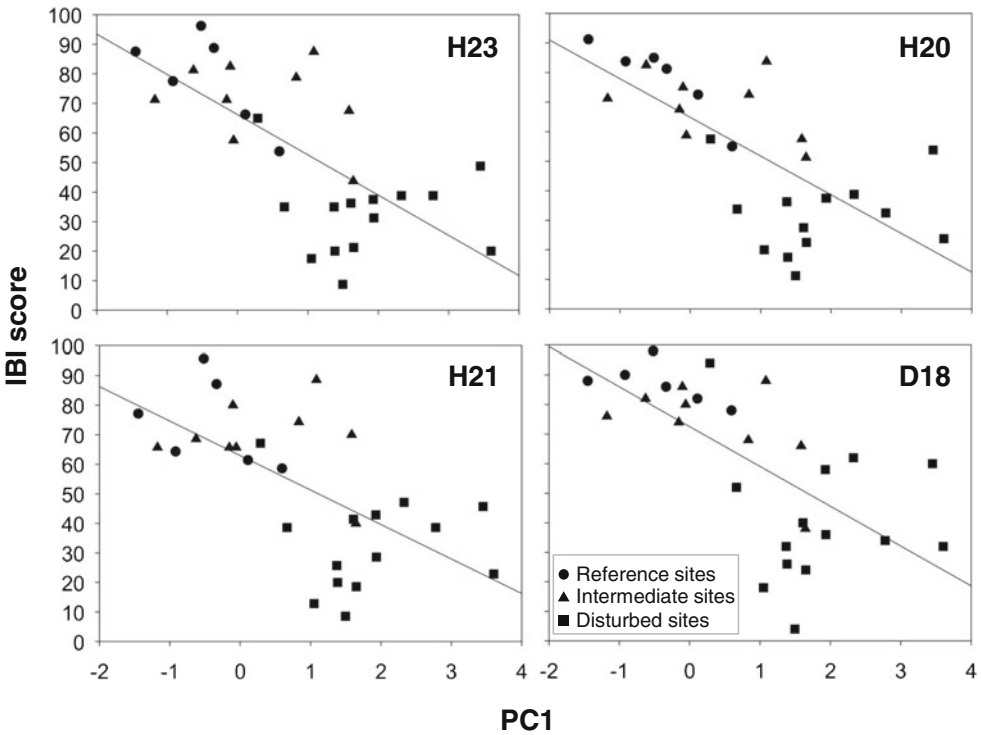

Fig. 2 Linear regression of IBI scores on water chemistry principal component scores. $P<0.0003$ for all relationships. Data are from southern California sites within the validation dataset. *Circles* reference sites, *triangles* Intermediate, *squares* Disturbed

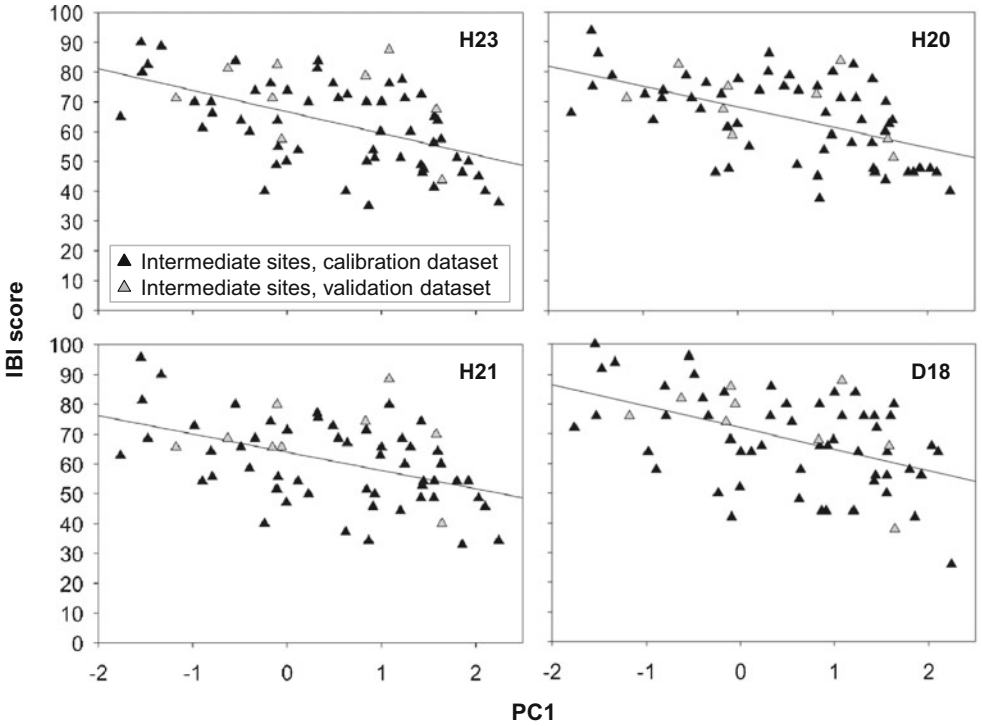

Fig. 3 Linear regression of IBI scores on water chemistry principal component scores. $P<0.0002$ for all relationships. Data are from the Intermediate site-disturbance class within southern California. *Gray* validation data, *black* calibration

values were 0.282 (H20), 0.273 (H23), 0.236 (D18), and 0.208 (H21) (Fig. 3). From the standpoint of proportion of variance explained, the top performer in each of the two analyses was the hybrid, H20.

As a measure of "repeatability," we also assessed how much IBI scores among replicate samples differed from one another by determining a "mean spread" value for each of the IBIs. Among replicates at each site, the lowest score for a given IBI was subtracted from the highest score for that same IBI. Then the mean of the resulting set of values across sites was calculated for each IBI, such that lower mean spread connoted higher repeatability for that IBI. Among field

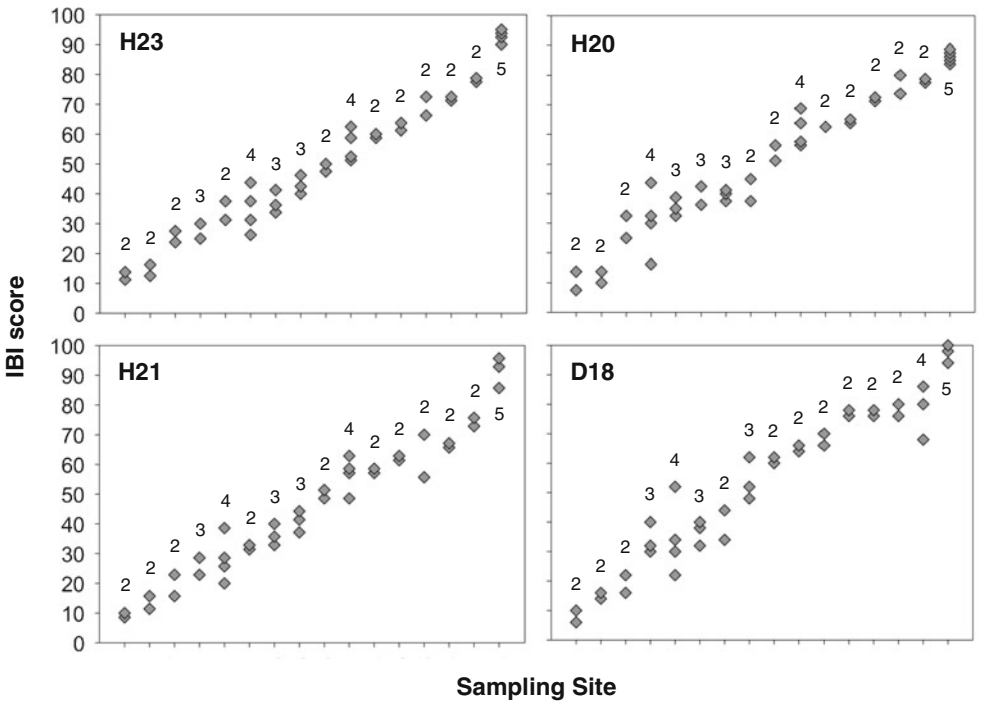

Fig. 4 Repeatability of IBI scores among replicate samples collected during a single site visit using the multihabitat field protocol at 16 sampling sites. Numbers associated with the data points for each site indicate the number of replicates collected at that site

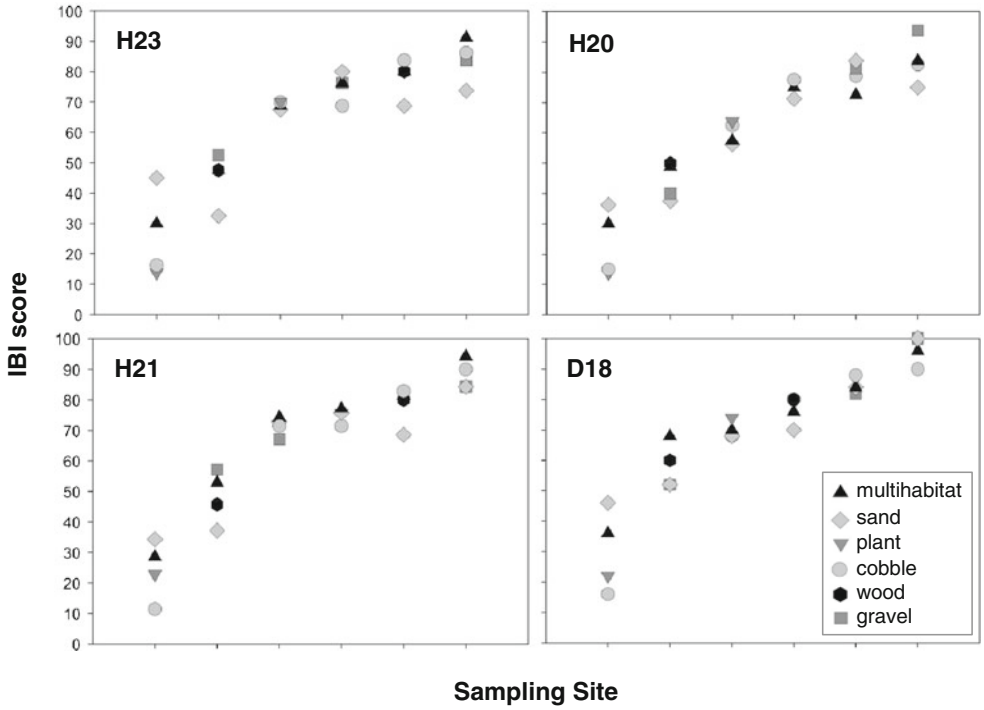

Fig. 5 "Repeatability" of IBI scores among samples collected from different substrata (N=4 samples/site) during a single site visit

replicates (from multihabitat sampling), mean spread varied from a score of 5 (on an overall IBI score scale of 0 to 100) corresponding to H23, to a mean spread of 6 for H20 and H21, and a mean spread of 8 for D18 (Fig. 4). For "repeatability" of IBI scores across different (targeted) substrate types, the broadest mean spread in scores, 16, corresponded to H23, and the other three IBIs exhibited a mean spread of 13 each (Fig. 5). No substratum type was associated with consistently highest or lowest scores for any of the IBIs, suggesting the absence of an effect of substratum on the scores of the IBIs tested.

Effects of varying levels of laboratory effort in soft algae analysis

To examine the potential effect of level of laboratory effort on IBI performance, three types of hybrid IBI were recognized: those including metrics that, in aggregate, require execution of the full soft algae laboratory protocol (Stancheva et al. 2012a), and those requiring reduced effort of two different types. The IBIs representing reduced effort were: H20 and H21. H21, which requires biovolume measurements but not species-level identification for soft algae, did not perform as well in the priority performance screens as H20, which relies on species information, but not biovolume. H20 and the top-performing, full-protocol hybrid (H23) were each superior with respect to different performance standards. H20 performed similarly to, or better than, H23 in terms of the various measures of responsiveness to stress, but H23 was superior in terms of signal-to-noise ratio and repeatability among replicate samples (albeit only when using the multihabitat sampling protocol), and H23 had a lower mean of pairwise correlations between metrics. However, none of the differences between H20 and H23, with the exception of signal-to-noise ratio, was particularly pronounced; hence our designation of H20 as overall top-performing IBI for southern California streams, when considering costs vs. benefits.

The boundary for IBI H20 score that we defined as that which distinguishes reference from non-reference sites based on statistical grounds, and which was calculated as two standard deviations below the mean IBI H20 score among the Reference sites in the project dataset, was determined to be 57.

Discussion

Approach to IBI development

Our approach to IBI development sought to ensure that metric selection was guided to the greatest extent possible by the species' known ecological traits. This is in contrast to an alternative approach: employing statistical modeling on the study dataset to reveal relationships between environmental variables and community measures as the basis for creating new metrics. Under this latter scenario, the strongest stressor–response relationships observed may not have obvious ecological underpinnings. Our approach had the dual advantages of (1) the indices being rooted in a priori knowledge of

species' ecological properties, such that they reflect the biological integrity of monitored streams, and (2) reducing the risk of "overfitting" the indices to the project dataset, which limits the degree to which they can be applied to new datasets (Hawkins 2004). That notwithstanding, it is worth noting that the boundary we propose for distinguishing sites from reference quality (i.e., an H20 score of 57) is based solely on statistical considerations, as opposed to specific knowledge that this score reflects an ecologically meaningful change point in community composition. As such, more work would be needed in order to establish the defensibility of 57 as an ecologically based threshold that could, for example, eventually be incorporated into a regulatory framework (e.g., to evaluate attainment of water body "aquatic life" goals).

Relative performance of IBIs

Although no IBI outperformed all others with respect to every criterion examined, from the standpoint of multiple performance criteria, hybrid IBIs were better than single-assemblage IBIs (Table 5). Our hybrid IBI scores, in general, corresponded more strongly to stressor gradients than single-assemblage IBIs. However, all of the top-performing IBIs across the different assemblage and effort categories exhibited responsiveness to stress; they all resulted in good separation of IBI scores between (at least) the Disturbed and Reference site classes.

Our results contrast the findings of Lavoie et al. (2004), who reported that incorporation of soft algae community composition information in ordination analyses did not improve upon diatoms' ability to distinguish among reference and agricultural streams, as well as the findings of Kelly et al. (2008), who reported that benthic soft algae did not improve predictability of chemical constituent concentrations in lakes (via the use of transfer functions), relative to benthic diatoms alone. These investigators used chiefly genus-level or coarser soft algae data in their analyses, which may account for some of the discrepancy with our findings. Indeed, among the hybrid IBI types that we explored, those requiring species-level soft algae data exhibited stronger stress response than the hybrid (H21) utilizing only coarser soft algae taxonomy.

Stress responses by the top-performing single-assemblage IBIs were grossly similar, but differed in that the soft algae IBI (S2) exhibited greater discriminatory power at the higher end of the range of site disturbance (i.e., between the Disturbed and Intermediate classes) than its diatom counterpart (D18). Within the Intermediate disturbance class, hybrid IBI scores were on average more strongly associated with the water-chemistry principal components axis than were single-assemblage IBI scores. IBI scores for sites with roughly comparable ecological condition may be more difficult to resolve within the Intermediate class than those for sites at the extreme ends of the disturbance gradient, due to the higher potential for a complex interplay of varying levels of multiple stressors within the intermediate-disturbance range. Our results suggest that use of hybrid indices that combine metrics from two assemblages that differ in terms of where, along a stressor gradient, they are most responsive may improve index responsiveness at intermediate levels of stress. This conclusion resonates with the findings of Schneider et al. (2013), who discovered differences in the responses of diatoms and soft algae communities in Norwegian streams to total phosphorus gradients (specifically, diatom taxa richness increased with total phosphorus, whereas soft algae richness decreased). Differences were also noted between the two assemblages in terms of where along a pH gradient taxa-richness values peaked. Schneider et al. (2013) concluded that relative influences of diatoms and soft algae on stream ecosystem structure and functioning vary according to certain factors (such as pH and nutrient supply), arguing for the inclusion of both assemblages in the assessment of phytobenthos structure and function.

As a group, hybrid IBIs exhibited much less metric redundancy (as measured by mean metric correlations) and marginally better signal-to-noise ratios (when using the multihabitat sampling protocol) than single-assemblage IBIs. Hybrid IBIs were also more indifferent to variation along natural gradients. Since diatom-based and soft algae-based IBIs were responsive to different sets of the natural-gradient types tested, the benefit of hybrids from this standpoint may lie in an "averaging" effect realized by mixing diatom metrics with soft algae metrics.

Addressing challenges associated with developing soft algae IBIs

In developing IBIs, we invested considerable efforts toward evaluating alternative approaches to using soft algae because, in general, less information on bioindicator development and performance is available for them than for diatoms. In addition to experimenting with different levels of laboratory effort, we examined the performance of data derived from different field-collection approaches, i.e., using only the purely objective, quantitative samples collected via the multihabitat method (for calculating biovolume-based metrics) vs. incorporating into the species-number metrics information from the epiphytes plus the qualitative samples, the latter of which were collected during sampling-reach macroalgal inventories. The inclusion of qualitative data in the species-number metrics helped to mitigate one of the challenges we encountered in using soft algae: specifically, sometimes only a low number of soft algae taxa was recorded from the quantitative sample at a given site, a phenomenon noted by Stevenson and Bahls (1999) as conferring an advantage to use of the comparatively species-rich diatoms over soft algae as bioindicators. Low soft algae species richness at a site could render the species-

number metrics more vulnerable to error. However, incorporating qualitative data in order to boost species numbers may increase susceptibility to sampling bias, assuming a higher potential for subjectivity associated with the collection of the qualitative sample. We accounted for this possibility by screening for and eliminating metrics that showed sensitivity to whether or not a qualitative sample had been collected. Nonetheless, any bioassessment program including qualitative data in an IBI should take measures to curb the potential for inconsistency among field crews through the administration of adequate training, intercalibration, and periodic auditing.

A second challenge presented by soft algae is that, for species-level identification, genera such as *Oedogonium*, *Mougeotia*, *Spirogyra*, *Zygnema*, and *Vaucheria*, require observation of reproductive structures, which are not always present on a given specimen (Kelly et al. 2008). We therefore chose to group specimens within these genera into loose taxonomic categories ("morphospecies"), based on readily determined morphological features such as filament width, number and type of chloroplasts, transverse cell wall type, and other vegetative characteristics (note that the designation of morphospecies occurred *prior* to embarking on IBI development). Some of the "taxa" with the highest indicator values resulting from indicator species analysis turned out to be morphospecies, which were therefore included as indicators in applicable metrics. In support of this practice, precedence exists for applying indicator values to soft algae morphosphecies for use in bioassessment (Schneider and Lindstrøm 2011).

A third challenge in using soft algae was the limited amount of published autecological data available for this assemblage, compared to diatoms. To compensate for this, we used indicator species analysis on our calibration dataset to establish taxon–stressor relationships upon which to base a small number of novel metrics for inclusion in applicable hybrid and soft algae-based IBIs. In order to root these metrics as much as possible in existing knowledge of algal ecological attributes, our choice of relationships to investigate was based on previous investigators' observations (see Methods) regarding environmental factors to which soft algae groups show sensitivity.

Our approaches to addressing the above challenges resulted in several high-performing IBIs incorporating soft algae metrics. We acknowledge that, on principle, inclusion of qualitative species-inventory data, use of morphospecies as indicator taxa, and/or development of metrics based on a subset of the project dataset may be objectionable to some. However, all three of these factors were at play in the case of IBI H20, which turned out to be our top performer in the validation exercise, suggesting that a robust IBI can be developed despite challenges inherent in using soft algae as a bioindicator.

Factors to weigh in selecting the optimal type of algae IBI for monitoring needs

Although we did not record the amount of taxonomy labor that was necessary for sample analysis, and the amount of time required to analyze a sample can vary widely as a function of factors like taxonomic diversity and number of uncommonly encountered species, we estimate that the amount of labor required for processing and analyzing a diatom sample using our protocol is on average roughly equal to the time required for processing and analyzing a soft algae sample when using either of our two, reduced-effort versions of Stancheva et al. (2012a). Furthermore, we estimate that conducting the full soft algae protocol adds 50 % laboratory labor to that which is required for a reduced-effort version (these estimates assume that analyses are carried out by experienced taxonomists, who are familiar with the regional flora). If laboratory effort relative to responsiveness to stress were the only two factors of concern in designing an algae assessment program, our results indicate that roughly 60 % better discriminatory power (in terms of variance in IBI scores explained by site-disturbance class) was realized for an additional 100 % laboratory effort, when comparing the top-performing reduced-effort hybrid (H20) with the top-performing diatom IBI (D18). Alternatively, roughly equal discriminatory power was realized for 25 % additional laboratory effort (overall) when comparing the top-performing full-effort hybrid IBI (H23) with H20. Cost vs. benefit is likely to be a major consideration for most monitoring programs in choosing which type of algae IBI to utilize. Table 5 provides a summary of our findings: IBI effort categories are ranked in terms of their relative performance (based on average results across IBIs within each category, from Table 4) for each performance criterion. Also shown is the relative amount of laboratory effort required for generating the data necessary for calculating each IBI type.

Both types of single-assemblage IBI performed reasonably well with respect to the priority performance criteria. Therefore, either assemblage (diatoms or soft algae) might be considered adequate for routine bioassessment. However, because more expertise is typically available for diatom analysis, they are likely to be the assemblage of choice for stream algal bioassessment if costs prohibit use of both assemblages. Alternatively, both assemblages might be used on a conditional basis. For instance, one might analyze diatom samples alone for basic screening assessments across a region, but analyze both assemblages for site-specific monitoring that requires higher resolution data (e.g., pursuant to a stream's nutrient total maximum daily load [TMDL] requirements). Alternatively, regardless of the application, a program might choose to analyze diatoms only, initially, at a given site, and invest in analysis of soft algae on a site-by-site basis, only when an ambiguous diatom IBI score is realized (e.g., to afford better discrimination among sites within the

intermediate range of site disturbance levels). Results of our analysis suggest that both alternatives would reduce net costs, yet allow the benefits of higher resolution information attainable from using both assemblages to be realized in situations in which that benefit will make the most difference.

IBI relative performance vs. cost is not the only factor to consider when choosing which assemblage(s) to monitor; other trade-offs also come into play. In addition the above-mentioned challenges of incorporating soft algae in bioassessment, there are also several potential advantages. (1) Diatoms tend to have high dispersal rates and short generation times, rendering them well suited to exhibiting rapid response to changes in their environment (Lavoie et al. 2008). Schneider et al. (2012) hypothesized soft algae to respond to environmental changes more slowly than diatoms, and Whitton (2012) noted that inclusion of relatively longer-lived soft algae taxa (e.g., *Batrachospermum*, *Lemanea*, and *Stigeoclonium*, and colonial species) in bioassessment efforts may result in better temporal integration of stress response than use of diatoms alone. (2) In terms of biomass, soft algae are often the major component of algae in a stream (Wehr and Sheath 2003) and most likely to manifest eutrophication in the form of nuisance blooms, arguing for documentation of the soft algae community for assessment of nutrient impacts. (3) Certain cyanobacterial taxa can produce cyanotoxins, which can negatively affect stream benthos. Data on soft algae taxonomic composition could thus be important for accurate interpretation of bioassessment data based on benthic macroinvertebrates (Aboal et al. 2002) and diatoms (Douterelo et al. 2004). (4) Some soft algae taxa (e.g., members of Rivulariaceae) respond to inorganic phosphate deficiency via development of long, colorless, multicellular hairs, which are the sites of phosphomonoesterase activity (Whitton and Mateo 2012), thus providing real-time diagnostic information about stream nutrient status.

Conclusion

Integrating information from two stream algal assemblages resulted in overall higher-performing IBIs than what we realized with diatoms or soft algae alone. Furthermore, our results indicated that an intermediate level of laboratory effort, specifically, one that forgoes soft algae biovolume information but maintains species-level taxonomy for that assemblage, yielded a hybrid IBI (H20) that is comparable to our top-performing full-effort hybrid (H23). Decision-making in designing a bioassessment program entails determining how to utilize the information obtained, and whether the value of results is cost-effectively enhanced through increasing levels of effort (Kelly 2006; Hughes and Peck 2008). It also requires consideration of other trade-offs inherent in using one vs. another assemblage. What magnitude of improved performance of a given IBI type merits its associated additional costs, and which of the other drawbacks and benefits inherent in the algal assemblage(s) comprising a candidate IBI matter the most, must ultimately be weighed by individual bioassessment programs.

Acknowledgments The California State Water Resources Control Board Consolidated Grants and Surface Water Ambient Monitoring Program (SWAMP) provided funding and data. Bérengère Laslandes, Christina Vanderwerken, Karen McLaughlin, Mariska Brady, Amanda Elliott, Evan Thomas, Andrew Fields, Liesl Tiefenthaler, and Nicholas Miller assisted in the field/laboratory. The Southern California Stormwater Monitoring Coalition and SWAMP-funded Reference Condition Management Program and Perennial Stream Assessment provided data. Rebecca Schaffner, Andrew Rehn, and Jason May provided landscape data and analyses. Marco Sigala provided database assistance. Martha Sutula, Julie Berkman, John Van Sickle, Scott Rollins, Yangdong Pan, Kenneth Schiff, Danuta Bennett, and Sarah Spaulding provided advice on study design/data analysis, and Stephen Weisberg, Joshua Westfall, Scott Johnson, and two anonymous reviewers provided valuable comments that improved the manuscript. Laboratories of the University of California, Santa Barbara, Marine Science Institute and the University of Georgia Odum School of Ecology performed water chemistry analyses.

References

Aboal M, Puig MA, Mateo P, Perona E (2002) Implications of cyanophyte toxicity on biological monitoring of calcareous streams in north-east Spain. J Appl Phycol 14:49–56

Bahls LL (1993) Periphyton bioassessment methods for Montana streams. Montana Water Quality Bureau, Department of Health and Environmental Science, Helena, USA

Barbour MT, Gerritsen J, White JS (1996) Development of the Stream Condition Index (SCI) for Florida. Florida Department of Environmental Protection, Tallahassee, USA

Cao Y, Hawkins CP, Olson J (2007) Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. J N Am Benth Soc 26:566–585

Cattaneo A, Kerimian T, Roberge M, Marty J (1997) Periphyton distribution and abundance on substrata of different size along a gradient of stream trophy. Hydrobiologia 354:101–110

Douterelo I, Perona E, Mateo P (2004) Use of cyanobacteria to assess water quality in running waters. Environ Pollut 127:377–384

Dufrêne M, Legendre P (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecol Monogr 67:345–366

Fernandez-Piñas F, Leganés F, Mateo P, Bonilla I (1991) Blue-green algae (cyanobacteria) as indicators of water quality in two Spanish rivers. In: Whitton BA, Rott E, Friedrich G (eds) Use of algae for monitoring rivers. Institut für Botanik, Universität Innsbruck, Innsbruck, pp 151–156

Fetscher AE, Busse LB, Ode PR (2009) Standard operating procedures for collecting stream algae samples and associated physical habitat and chemical data for ambient bioassessments in California. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 002

Fore LS (2003) Developing biological indicators: Lessons learned from mid-Atlantic streams. EPA 903/R-003/003. U.S. Environmental Protection Agency, Office of Environmental Information and Mid-Atlantic Integrated Assessment Program, Region 3, Ft. Meade, USA

Fore L, Grafe C (2002) Using diatoms to assess the biological condition of large rivers in Idaho (U.S.A.). Freshw Biol 47:2015–2037

Griffith MB, Hill BH, Herlihy AT, Kaufmann PR (2002) Multivariate analysis of periphyton assemblages in relation to environmental gradients in Colorado Rocky Mountain streams. J Phycol 38:83–95

Guasch H, Paulsson M, Sabater S (2002) Effect of copper on algal communities from oligotrophic calcareous streams. J Phycol 38:241–248

Gutowski A, Foerster J, Schaumburg J (2004) The use of benthic algae, excluding diatoms and Charales, for the assessment of the ecological status of running waters: a case history from Germany. Oceanol Hydrobiol Stud 33:3–15

Hawkins DM (2004) The problem of overfitting. J Chem Inf Comput Sci 44:1–12

Hering D, Feld CK, Moog O, Ofenböck T (2006) Cook book for the development of a multimetric index for biological condition of aquatic ecosystems: experiences from the European AQEM and STAR projects and related initiatives. Hydrobiologia 566:311–324

Hill BH, Herlihy AT, Kaufmann PR, Stevenson RJ, McCormick FH, Burch Johnson C (2000) Use of periphyton assemblage data as an index of biotic integrity. J N Am Benth Soc 19:50–67

Hill BH, Stevenson RJ, Pan Y, Herlihy AT, Kaufmann PR, Burch Johnson C (2001) Comparison of correlations between environmental characteristics and stream diatom assemblages characterized at genus and species levels. J N Am Benth Soc 20:299–310

Hughes RM, Peck DV (2008) Acquiring data for large aquatic resource surveys: the art of compromise among science, logistics, and reality. J N Am Benth Soc 27:837–859

John MD (2011) Order Chaetophorales, Microsporales, Ulotrichales. In: John DM, Whitton BA, Brook AJ (eds) The freshwater algal flora of the British Isles. An identification guide to freshwater and terrestrial algae. Cambridge University Press, Cambridge, pp 524–553

John DM, Johnson LR (1991) Green microphytic algae as river water quality monitors. In: Whitton BA, Rott E, Friedrich G (eds) Use of algae for monitoring rivers. Institut für Botanik, Universität Innsbruck, Innsbruck, pp 151–156

Kaufmann PR, Levine P, Robison EG, Seeliger C, Peck DV (1999) Quantifying physical habitat in wadeable streams. EPA/620/R-99/003. U.S. Environmental Protection Agency, Washington

Kelly MG (2006) A comparison of diatoms with other phytobenthos as indicators of ecological status in streams in northern England. In: Witkowski A (ed) Proceedings of the 18th International Diatom Symposium. Biopress, Bristol, pp 139–151

Kelly MG, King L, Jones RI, Barker PA, Jamieson BJ (2008) Validation of diatoms as proxies for phytobenthos when assessing ecological status in lakes. Hydrobiologia 610:125–129

Klemm DJ, Blocksom KA, Thoeny WT, Fulk FA, Herlihy AT, Kaufmann PR, Cormier SM (2002) Methods development and use of macroinvertebrates as indicators of ecological conditions for streams in the Mid-Atlantic Highlands region. Env Mon Assess 78:169–212

Komárek J, Kling H, Komárková J (2002) Filamentous cyanobacteria. In: Wehr JH, Sheath RG (eds) Freshwater algae of North America. Ecology and classification. Academic Press, San Diego, pp 59–191

Lavoie I, Vincent WF, Pienitz R, Painchaud J (2004) Benthic algae as bioindicators of agricultural pollution in the streams and rivers of southern Québec (Canada). Aquat Ecosyst Health Manage 7:43–58

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

Lavoie I, Campeau S, Darchambeau F, Cabana G, Dillon PJ (2008) Are diatoms good integrators of temporal variability in stream water quality? Freshw Biol 53:827–841

Lavoie I, Dillon PJ, Campeau S (2009) The effect of excluding diatom taxa and reducing taxonomic resolution on multivariate analyses and stream bioassessment. Ecol Ind 9:213–225

Leland HV, Porter SD (2000) Distribution of benthic algae in the upper Illinois River basin in relation to geology and land use. Freshw Biol 44:279–301

Lowe RL, Laliberte GD (1996) Benthic stream algae: distribution and structure. In: Hauer FR, Lamberti GA (eds) Methods in stream ecology. Academic Press, San Diego, pp 269–293

McCune B, Grace JB (2002) Analysis of ecological communities. MjM Software, Gleneden Beach

Moulton SR, Kennen JG, Goldstein RM, Hambrook JA (2002) Revised protocols for sampling algal, invertebrate, and fish communities as part of the National Water-Quality Assessment Program. USGS Open-File Report 02–150. U.S. Geological Survey, Reston

Ode PR, Rehn AC, May JT (2005) A quantitative tool for assessing the integrity of southern coastal California streams. Env Manag 35:493–504

Palmer CM (1969) A composite rating of algae tolerating organic pollution. J Phycol 5:78–82

Perona E, Bonilla I, Mateo P (1998) Epilithic cyanobacterial communities and water quality: an alternative tool for monitoring eutrophication in the Alberche River (Spain). J Appl Phycol 10:183–191

Pipp E, Rott E (1996) Recent developments in the use of benthic algae (excluding diatoms) for monitoring rivers in Austria and Germany. In: Whitton BA, Rott E (eds) Use of algae for monitoring rivers II. Institut für Botanik, Universität Innsbruck, Innsbruck, pp 160–165

Porter SD, Mueller DK, Spahr NE, Munn MD, Dubrovsky NM (2008) Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. Freshw Biol 53:1036–1054

Potapova M, Charles DF (2007) Diatom metrics for monitoring eutrophication in rivers of the United States. Ecol Indic 7:48–70

Power ME (1990) Effects of fish in river food webs. Science 250:811–814

PRISM Climate Group (2013) Oregon State University, http://prism.oregonstate.edu. Accessed 13 April 2013

Rott E, Hofmann G, Pall K, Pfister P, Pipp E (1997) Indikationslisten für Aufwuchsalgen in Fließgewässern in Österreich: Teil 1. Saprobielle Indication. Projekt des Bundesministeriums für Land- und Forstwirtschaft,Wasserwirtschaftskataster, f. 1–80

Rott E, Pipp E, Pfister P, Van Dam H, Ortler K, Binder N, Pall K (1999) Indikationslisten für Aufwuchsalgen in Österreichischen Fließgewässern: Teil 2. Trophieindication. Bundesministerium f. Land- und Forstwirtschaft, Zahl 41.034/08- IVA 1/97, Wien. f. 1–248

Rusanov AG, Stanislavskaya E, Ács E (2012) Periphytic algal assemblages along environmental gradients in the rivers of the Lake Ladoga basin, northwestern Russia: implication for the water quality assessment. Hydrobiologia 695:305–327

Schaumburg J, Schranz C, Foerster J, Gutowski A, Hofmann G, Meilinger P, Schneider S, Schmedtje U (2004) Ecological classification of macrophytes and phytobenthos for rivers in Germany according to the Water Framework Directive. Limnologica 34:283–301

Schneider S (2011) Impact of calcium and TOC on biological acidification assessment in Norwegian rivers. Sci Total Env 409:1164–1171

Schneider SC, Lindstrøm E-A (2009) Bioindication in Norwegian rivers using non-diatomaceous benthic algae: the acidification index periphyton (AIP). Ecol Indic 9:1206–1211

Schneider SC, Lindstrøm E-A (2011) The periphyton index of trophic status PIT: a new eutrophication metric based on non-diatomaceous benthic algae in Nordic rivers. Hydrobiologia 665:143–155

Schneider SC, Lawniczak AE, Picińska-Faltynowicz J, Szoszkiewicz K (2012) Do macrophytes, diatoms and non-diatom benthic algae give redundant information? Results from a case study in Poland. Limnologica 42:204–211

Schneider SC, Kahlert M, Kelly MG (2013) Interactions between pH and nutrients on benthic algae in streams and consequences for

ecological status assessment and species richness patterns. Sci Total Env 444:73–84

Sheath RG (2003) Red algae. In: Wehr JD, Sheath RG (eds) Freshwater algae of North America: ecology and classification. Academic Press, San Diego, pp 197–221

Sheath RG, Cole KM (1992) Biogeography of stream macroalgae in North America. J Phycol 28:448–460

Sheath RG, Morison MO, Korch JE, Kaczmarczyk D, Cole KM (1986) Distribution of stream macroalgae in south-central Alaska. Hydrobiologia 135:259–269

Sládeček V (1973) System of water quality from the biological point of view. Arch Hydrobiol Beih Ergeb Limnol 7:1–218

Spaulding SA, Lubinski DJ, Potapova M (2010) Diatoms of the United States. http://westerndiatoms.colorado.edu. Accessed 10 April 2013

Stancheva R, Fetscher AE, Sheath RG (2012a) A novel quantification method for stream-inhabiting, non-diatom benthic algae, and its application in bioassessment. Hydrobiologia 684:225–239

Stancheva R, Hall JD, Sheath RG (2012b) Systematics of the genus *Zygnema* (Zygnematophyceae, Charophyta) from Californian watersheds. J Phycol 48:409–422

Stevenson RJ, Bahls LL (1999) Periphyton protocols. In: Barbour MT, Gerritsen J, Snyder BD (eds) Rapid bioassessment protocols for use in wadeable streams and rivers: periphyton, benthic macroinvertebrates, and fish, EPA 841-B-99-002. US Environmental Protection Agency, Washington, pp 6-1–6-22

Stevenson RJ, Pan Y, van Dam H (2010) Assessing environmental conditions in rivers and streams with diatoms. In: Smol JP, Stoermer EF (eds) The diatoms: applications for the environmental and earth sciences, 2nd edn. Cambridge Univ. Press, Cambridge, pp 57–85

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

Stoddard JL, Herlihy AT, Peck DV, Hughes RM, Whittier TR, Tarquinio E (2008) A process for creating multimetric indices for large-scale aquatic surveys. J N Am Benth Soc 27:878–891

van Dam H, Mertens A, Sinkeldam J (1994) A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. Neth J Aquat Ecol 28:117–133

Van Der Werff A (1955) A new method of concentrating and cleaning diatoms and other organisms. Proc Int Assoc Theor Appl Limnol 12:276–277

Van Sickle J (2010) Correlated metrics yield multimetric indices with inferior performance. Trans Am Fisher Soc 139:1802–1817

VanLandingham SL (1982) Guide to the identification, environmental requirements and pollution tolerance of bluegreen algae (Cyanophyta). EPA-600/3-82-073

Vis C, Hudon C, Cattaneo A, Pinel-Alloul B (1998) Periphyton as an indicator of water quality in the St. Lawrence River (Québec, Canada). Env Pollut 101:13–24

Wang Y-K, Stevenson RJ (2005) Development and evaluation of a diatom-based index of biotic integrity for the Interior Plateau Region, USA. J N Am Benth Soc 24:990–1008

Wehr JD, Sheath RG (2003) Freshwater habitats of algae. In: Wehr JD, Sheath RG (eds) Freshwater algae of North America: ecology and classification. Academic Press, San Diego, pp 11–57

Whitton BA (2012) Changing approaches to monitoring during the period of the 'Use of Algae for Monitoring Rivers' symposia. Hydrobiologia 695:7–16

Whitton BA, Mateo P (2012) Rivulariaceae. In: Whitton BA (ed) Ecology of cyanobacteria: II. Their diversity in space and time. Springer, Dordrecht, pp 561–591