J. Phycol. 44, 787–802 (2008)
© 2008 Phycological Society of America
DOI: 10.1111/j.1529-8817.2008.00523.x

COMPARISON OF SIMPLE AND MULTIMETRIC DIATOM-BASED INDICES FOR GREAT LAKES COASTLINE DISTURBANCE[1]

Euan D. Reavie[2], *Amy R. Kireta, John C. Kingston*

Center for Water and the Environment, Natural Resources Research Institute, University of Minnesota Duluth, 1900 East Camp Street, Ely, Minnesota 55731, USA

Gerald V. Sgro

Biology Department, John Carroll University, 20700 North Park Blvd., University Heights, Ohio 44118, USA

Nicholas P. Danz, Richard P. Axler, and Thomas P. Hollenhorst

Center for Water and the Environment, Natural Resources Research Institute, University of Minnesota Duluth, 5013 Miller Trunk Highway, Duluth, Minnesota 55811, USA

Because diatom communities are subject to the prevailing water quality in the Great Lakes coastal environment, diatom-based indices can be used to support coastal-monitoring programs and paleoecological studies. Diatom samples were collected from Great Lakes coastal wetlands, embayments, and high-energy sites (155 sites), and assemblages were characterized to the species level. We defined 42 metrics on the basis of autecological and functional properties of species assemblages, including species diversity, motile species, planktonic species, proportion dominant taxon, taxonomic metrics (e.g., proportion Stephanodiscoid taxa), and diatom-inferred (DI) water quality (e.g., DI chloride [Cl]). Redundant metrics were eliminated, and a diatom-based multimetric index (MMDI) to infer coastline disturbance was developed. Anthropogenic stresses in adjacent coastal watersheds were characterized using geographic information system (GIS) data related to agricultural and urban land cover and atmospheric deposition. Fourteen independent diatom metrics had significant regressions with watershed stressor data; these metrics were selected for inclusion in the MMDI. The final MMDI was developed as the weighted sum of the selected metric scores with weights based on a metric's ability to reflect anthropogenic stressors in the adjacent watersheds. Despite careful development of the multimetric approach, verification using a test set of sites indicated that the MMDI was not able to predict watershed stressors better than some of the component metrics. From this investigation, it was determined that simpler, more traditional diatom-based metrics (e.g., DI Cl, proportion Cl-tolerant species, and DI total phosphorus [TP]) provide superior prediction of overall stressor influence at coastal locales.

[1]Received 14 May 2007. Accepted 27 November 2007.
[2]Author for correspondence: e-mail ereavie@nrri.umn.edu.

Key index words: **coastlines; diatoms; eutrophication; Great Lakes; indices; metrics; multimetric; stressors**

Abbreviations: **1/TTube, inverted transparency tube measurement; AG, agricultural principal component; ATM, atmospheric principal component; Cl, chloride; DCA, detrended correspondence analysis; DI, diatom-inferred; EBF, eastern broadleaf forest; GIS, geographic information system; GLEI, Great Lakes Environmental Indicators; IND, industrial principal component; LMF, Laurentian mixed forest; MAXREL, maximum relative stressor value; MMDI, multimetric diatom index; PCA, principal components analysis; SUMREL, standardized composite stressor value; TP, total phosphorus; TSS, total suspended solids; URB, urban development principal component; WA, weighted average**

The condition of the Laurentian Great Lakes coastal environments has received much attention in recent years (Keough and Griffin 1994, Maynard and Wilcox 1997, Lawson 2004, Brazner et al. 2007), but no comprehensive long-term strategy is in place to assess the condition of these environments and to monitor environmental impacts of human activities. Biological indicators of coastal water quality have become a mainstay of ecological assessments because they reflect the impacts of watershed activities on adjacent aquatic environments and have advantages over physicochemical methods (Hellawell 1986, Reavie et al. 2006). Many water resource management strategies now rely on biotic indices (i.e., assessments based on resident floral and faunal communities). Although biotic index approaches in the Great Lakes have gained attention in the last decade (Wilcox et al. 2002, Environment Canada & U.S. EPA 2003, Albert and Minc 2004, Niemi and McDonald 2004, Uzarski et al. 2004), most indicators proposed for Great Lakes coastal environments remain uncalibrated and untested throughout large parts of the basin or under the full range of environmental conditions. Developing effective indicators of ecological condition requires that indicators be calibrated to identify the accuracy and precision of their responses to important environmental stressors (Karr and Chu 1999, Seegert 2001) to quantify their value in monitoring and assessment programs.

Organism-based indices have been a valuable monitoring tool in lotic waters and small lake systems. Index approaches have been commonly applied for fish (Karr et al. 1986) and invertebrate (Kerans and Karr 1994) indices, and algal indices (Rosen 1995, Whitton and Kelley 1995, Chessman et al. 1999, Hill et al. 2000, Fore and Grafe 2002) are now being incorporated into routine monitoring assessments in the U.S. (Charles 1996, Leland and Porter 2000). We expect that similar index approaches will be a valuable addition to monitoring programs in the Great Lakes.

Many studies have linked changes in algal assemblages, particularly diatoms, to changes in water chemistry, such as pH, nutrients, and salinity (Smol 2002), with the main goal being to identify and validate environmental optima and tolerances of indicator species. This approach is understandable because water chemistry variables are meaningful proxies of human disturbance (e.g., the inevitable increase in nutrient concentrations resulting from agricultural practices). Algal indicators also have the potential to provide an integrated assessment by evaluating indirect algal responses to human disturbances, such as adjacent agricultural and urban development (Chessman et al. 1999, Hill et al. 2000). Diatoms have become the most widely applied indicator group because they are taxonomically distinct, abundant in almost all aquatic environments, respond rapidly to changing conditions, and are well preserved in sediment deposits (Hall and Smol 1999). Researchers can use the percentages of certain diatom taxa to classify and quantify long-term environmental changes that result from anthropogenic activities. Correlations between DI water quality and watershed characteristics, such as urban and agricultural extent, have provided an important link between bioindicators and anthropogenic influences in watersheds (Dixit and Smol 1994). Furthermore, evaluations of diatoms can provide a description of water quality that is not achievable from snapshot chemical analyses; the value of an integrative biological response can offset the inconsistency of rapid changes in water chemistry (Reavie et al. 2006).

Paleolimnologists have developed and applied robust diatom-based models. Such models are typically constructed by assuming unimodal responses of the species across environmental gradients and calibrating diatom species' responses to measured water quality variables in large lake sets. These models appeal to water quality managers because they provide inferred quantitative data for specific variables such as nutrients. Such a model was recently developed for Great Lakes coastlines (Reavie et al. 2006). While this Great Lakes model will be of interest to managers and paleoecologists, it has some logistic constraints (e.g., time and monetary dedication, taxonomic expertise, specialized software, long learning curve) that may not make it the best choice for all managers who may be considering the algae as an environmental quality indicator. Index approaches provide a means to evaluate environmental quality at a locale based on the diatom assemblage and can be flexible enough to minimize some of the aforementioned constraints and suit a greater user audience. For instance, an index can be calibrated using genera- instead of species-level taxonomic assessment, or by assessing functional groups of taxa, thus reducing sample assessment effort. Furthermore, algal indices can simultaneously include several characteristics of the assemblage at a locale, potentially providing an integrated picture of impacts at a site by not being limited to inferring a specific water quality parameter.

Several candidate metrics were derived from data collected as part of a larger study designed to develop and test indicators of ecological condition for Great Lakes coastal ecosystems (the Great Lakes Environmental Indicators [GLEI] project, http://glei.nrri.umn.edu; Niemi et al. 2004, Danz et al. 2005). The larger study included collection of abundance information on other biotic assemblages, including birds, fish, amphibians, aquatic macroinvertebrates, and wetland vegetation from the U.S. coastal locations spread throughout each of the Great Lakes.

Useful indices of environmental quality are responsive to anthropogenic stressors and have well-understood, unidirectional responses. While it is known that metrics, such as algal species diversity, can reflect limnological condition (Patrick 1973), rigorous testing of potential metrics is recommended before they are applied. It is noteworthy that two European indices, the Lange-Bertalot Index (Lange-Bertalot 1979) and the Trophic Diatom Index (Kelly and Whitton 1995), have been applied to these Great Lakes data (Brazner et al. 2007). Although these two metrics had some capacity to track human impacts in Great Lakes coastlines, for this study, it was our desire to use metrics that focused on local Great Lakes data and required a minimum of additional reference material.

We aimed to create an MMDI on the basis of the ''best'' of the candidate metrics. Simply combining multiple metrics into a single multimetric index is problematic because it assumes that each metric is independent and of equal importance. We eliminated metrics that provided redundant information about environmental quality and only included metrics that had significant observed relationships with anthropogenic stressors. Metrics reflected taxonomic (e.g., proportion of a particular genus) and functional (e.g., proportion with a particular adaptive strategy, such as the ability to assimilate atmospheric nitrogen) characteristics of the assemblage. These metrics were regressed against watershed characteristics, such as agricultural intensity and urban development, to identify the power of each metric to reflect anthropogenic stress. Our objective was to develop a diatom-based index of coastal disturbance similar to those developed for fish (Karr et al. 1986), macroinvertebrates (Kerans and Karr 1994), and periphyton (Hill et al. 2000), which could be used for the biological assessment of Great Lakes coastal ecosystems. Validation included comparing MMDI scores to watershed stressor data and determining if the MMDI performed better than the best of the component metrics and diatom-based index approaches that are already available.

MATERIALS AND METHODS

Sampling design. Coastal sample locations were selected as described by Danz et al. (2005). Briefly, the entire U.S. coastline of the Great Lakes was broken up into 762 segments, each consisting of a shoreline reach (''segment'') and associated watershed (''segment-shed''). Each segment-shed was summarized using 207 GIS-based environmental variables that included anthropogenic activities and soil data. Cluster analysis was carried out using these environmental data to create groups of segment-sheds with similar environmental profiles. Sample locations were randomly selected from each cluster, excluding inaccessible locations. This selection method provided a subset of sample locations that reflected the range of natural and anthropogenic environmental conditions present along the Great Lakes shorelines.

Five coastal ecosystem types, as described by Danz et al. (2005), were sampled (Fig. 1): embayments, high-energy shorelines, coastal wetlands, riverine wetlands, and protected wetlands. A total of 155 coastal sites had sufficiently complete diatom and environmental data for consideration in this study. A related investigation (Kireta et al. 2007) determined that ecosystem type has some influence on the structure of Great Lakes diatom assemblages, but that evaluating diatom-environmental relationships on the combined dataset of ecosystem types provided the best autecological information.

Field work, sample preparation, and diatom analysis. Field sites were sampled from June to September 2002 and May to August 2003. A detailed suite of environmental measurements was collected at each sample location, and full sampling and analytical methods are provided by Reavie et al. (2006). Quality assurance/quality control procedures followed a Quality Assurance Project Plan submitted to the U.S. Environmental Protection Agency (EPA) at the start of the project that followed EPA guidelines (U.S. EPA 1999).

Benthic and sedimented diatoms were sampled from natural substrates from 0.5 to 3 m water depths. Surface sediments were preferentially sampled, but 12% of sites required the sampling of epilithon due to unsuitable sediment regimes. Sampling and preparation methods for diatoms are also described by Reavie et al. (2006). Diatom assemblages for each sample were represented by at least 400 valves counted along slide transects at ×1,000 magnification using oil immersion microscopy. Individual valves were identified to the lowest taxonomic level possible using numerous diatom checklists and iconographs (see Reavie et al. 2006).

Statistical analyses. Forty-two candidate metrics were tested to identify their potential as components of a multimetric diatom index (MMDI; Table 1). Metrics were compared to two main parameters: ''stress'' and the ''natural gradient.'' In this article, stress (or stressors) refers to anthropogenic factors that result in possible or realized impacts on limnological quality (e.g., agriculture, urban development), whereas the natural gradient refers to ranges of factors that have little or no anthropogenic component (e.g., spatial coordinate, water temperature). We assumed that ideal metrics for multimetric compilation are those that were (A) correlated to gradients of anthropogenic disturbance and (B) uncorrelated with other metrics. Some metrics, such as DI TP, were expected to be strongly correlated to environmental variables that reflect anthropogenic stress, such as the proportion of agricultural land cover in a watershed. Other metrics, such as the relative abundance of *Planothidium* species, could not be predicted because the environmental affinities (in terms of anthropogenic stressors) of these metrics have not yet been clearly defined in the literature. The candidate metrics we tested are described below.

DI water quality variables: Diatom transfer functions were derived by relating diatom species assemblages in the coastal samples to corresponding measured water quality variables (Reavie et al. 2006). The resulting transfer function for each variable comprises species coefficients (i.e., environmental optima) that can be used to infer quantitative information about the variable, based on the relative abundance of each species in a given sample. Particularly robust diatom-based models were derived for five variables that are important water quality indicators: TP, a critical nutrient that is directly related to environmental problems such as cultural eutrophication (Schindler 1977); chl *a*, a direct indicator of algal standing crop and a useful indicator of nutrient load and consequent algal growth; total suspended solids (TSS), a proxy for inorganic particulate load and water clarity; transparency tube measurement (1/TTube), a simple method for estimating transparency, inverted to make the data directly comparable to turbidity measurements (Anderson and Davic 2004); and chloride (Cl), a potential indicator of pollution from road salt applications and other sources (Godwin et al. 2003). Additional variables (e.g., suspended solids and nitrogen compounds) also produced robust diatom-based models, but these variables strongly covaried with the selected variables and so were predetermined as redundant.

Full details of the development of diatom inference models are provided in the associated document (Reavie et al. 2006), but the abbreviated methods are as follows: Transfer functions were developed using weighted averaging (WA) regression and calibration (C2 software; Juggins 2003). DI estimates of water quality variables for each sample were calculated by taking the optimum of each species to that variable, weighting it by its abundance in that sample, and calculating the average of the combined weighted species optima. Because of previous analyses, we were already aware that DI water quality data were correlated with watershed stressor data; however, these relationships had not yet been considered in the context of other potential metrics. We expected the five selected DI parameters to be positively correlated with environmental disturbance.

Shannon–Weaver index of diversity: Metrics, such as species richness and evenness, are often used to describe species distributions, but they rely on fixed sample areas or volumes and complete assessment of species composition in a sample to derive values. These are not practical candidate metrics, because we assessed only portions of our samples (i.e., microscope slide transects), and good estimates of diatom species richness at a site would require counts into the

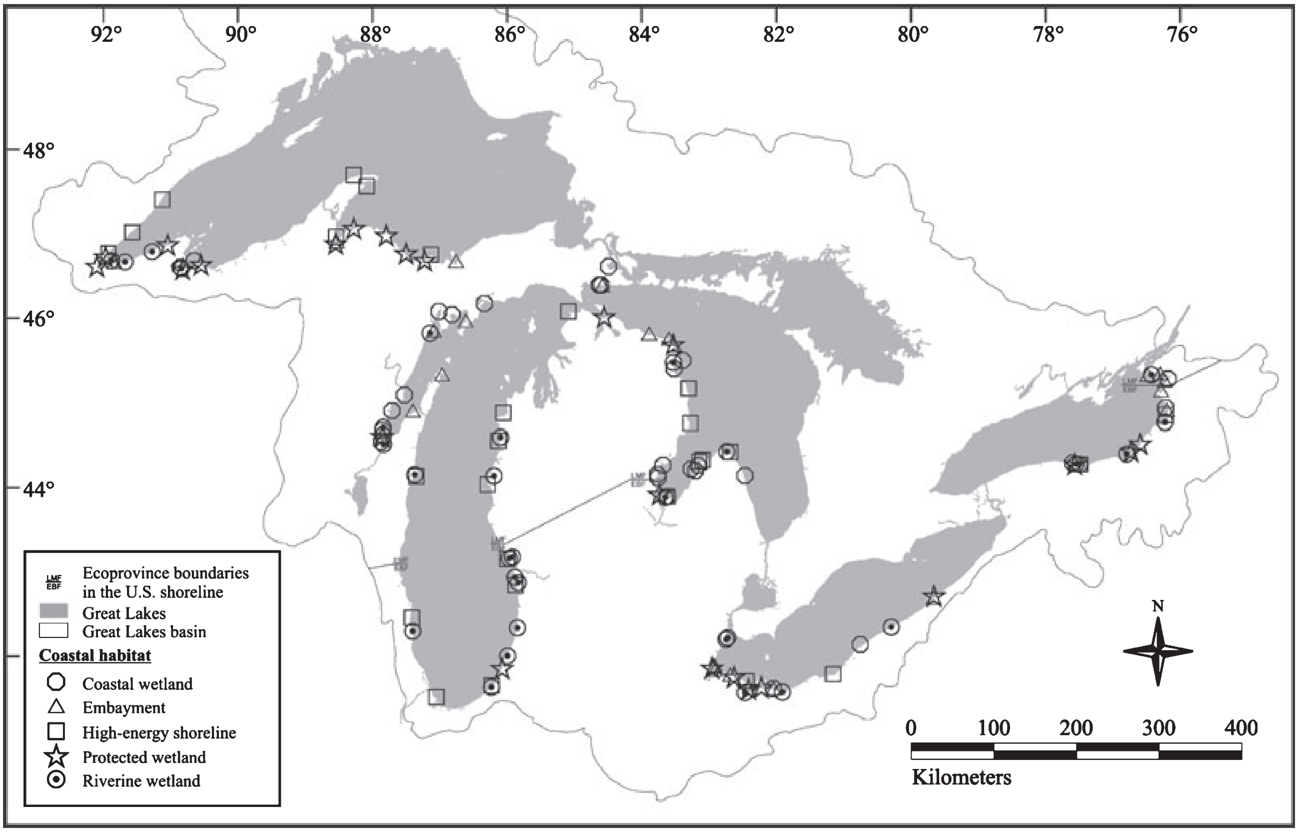

FIG. 1. Location map of the Great Lakes study area.

thousands or higher to identify very rare taxa (Patrick et al. 1954). Instead, we relied on the diversity of the assemblage data from each sample because diversity is likely to be highly correlated to richness and evenness. Two candidate diversity metrics were calculated: one based on the highest taxonomic resolution (>2,000 taxa) and another based on the genera (98 genera). It was expected that these two metrics would strongly covary, and so at least one of them would be eliminated from multimetric consideration. Both metrics were tested to determine if lower taxonomic resolution (genus) could provide meaningful environmental information and a potential means for rapid assessment of the diatom assemblage.

Diatom species diversity usually decreases as a result of disturbance such as cultural eutrophication (Williams 1964, Lotter 2001), although some researchers have reported increases in diatom diversity under moderate stress (Stevenson 1984). Despite conflicting observations in the literature, on the basis of cursory analyses of the diatom data, we expected diversity to be inversely related to environmental disturbance.

DCA score: Detrended correspondence analysis (DCA), with detrending by segments and downweighting of rare taxa, was used to define the major gradient of floristic variation in the diatom data (Hill and Gauch 1980). DCA assigns axis scores on the basis of assemblage similarity, with the first axis capturing the most important gradient of variation in the diatom data. For example, two samples with similar diatom assemblages will have similar DCA scores, whereas the two most dissimilar assemblages will likely lie on opposite ends of the DCA gradient. The diatom-based DCA score has previously been used successfully to develop a disturbance index for the Environmental Monitoring and Assessment Program-Surface Waters (EMAP-SW; Dixit and Smol 1994). Although a strong correlation between environmental disturbance and DCA score was anticipated, the direction of this correlation could not be predicted, because the ''low'' and ''high'' ends of the DCA gradient are arbitrary. Because the first two DCA axes reflect the two most important gradients capturing variation in the assemblage data, these two axes (i.e., their sample score data sets) were evaluated as candidate metrics.

The use of the DCA score as a metric may be somewhat less accessible to some users because the DCA score of a new diatom sample can only be derived by treating the new sample as supplementary in a DCA with the ''active'' diatom samples that were used to derive the index (i.e., the entire suite of Great Lakes diatom data). Provided an analyst has the original calibration data set and appropriate software, such as CANOCO (Ter Braak and Šmilauer 2002), this is a rapid calculation.

Proportion of the most dominant diatom: The relative abundance of the dominant diatom taxon in an assemblage can reflect the evenness of the taxonomic representation in an assemblage. Species that are better adapted to stressful conditions, such as nutrient enrichment, will have a competitive advantage during an increase in nutrient flux, resulting in an uneven distribution of individuals among taxa (Gray 1989). Often, dominance by a single *r*-selected taxon is noted following physical or chemical disturbance, or persistently in areas where disturbance events are frequent. On the basis of preliminary analysis, we noted that natural physical disturbance had an influence on the dominant diatom metric. High-energy and embayment locales were more likely to have higher values of the dominant diatom metric than wetland locales (ANOVA; $P < 0.0001$). Here, we further evaluated this metric in terms of anthropogenic stressors. We anticipated that the percent

TABLE 1. Candidate metrics characterized for the Great Lakes coastal locations, and the decision matrix to select metrics for inclusion in multimetric development, based on the ratio of anthropogenic to natural response.

Metric	Minimum	Maximum	Average	Transformation applied	Anticipated response to disturbance	Actual response to disturbance	$R^2_{anthropogenic}$	$R^2_{natural}$	$R^2_{anthropogenic}$: $R^2_{natural}$ ratio	Cluster	Decision
DI TP (log [μg · L^{-1} + 1])	0.61	2.35	1.43	None	↑	AG(↑), ATM(↑)	0.28	0.66	0.42	3	Reject
DI chl *a* (log [ppb + 1])	0.00	1.80	0.72	None	↑	AG(↑), IND(↑)	0.25	0.62	0.41	3	Reject
DI TSS (log [mg · L^{-1} + 1])	0.10	1.90	0.90	None	↑	AG(↑), IND(↑)	0.24	0.59	0.40	3	Reject
DI 1/TTube (log [cm^{-1}] × 10^{-2})	0.05	3.15	0.92	None	↑	AG(↑), ATM(↑), IND(↑)	0.22	0.60	0.37	3	Reject
DI Cl (log [mg · L^{-1} + 1])	0.16	1.87	1.15	None	↑	AG(↑), ATM(↑), IND(↑), URB(↑)	0.48	0.72	0.66	3	Accept
Shannon–Weaver (total)	1.39	4.36	3.21	None	↓		0.06	0.24	0.24	4	Reject
Shannon–Weaver (genus)	1.10	3.03	2.25	None	↓		0.05	0.19	0.25	4	Reject
DCA AX1	0.00	4.27	1.57	None	?	ATM(↑), IND(↑), URB(↑)	0.28	0.63	0.43	3	Reject
DCA AX2	0.00	4.01	2.20	None	?		0.03	0.13	0.24	5	Reject
% Dominant diatom	6.31	68.95	20.67	log(x + 1)	↑		0.03	0.23	0.12	4	Reject
% Motile	2.76	88.89	34.88	None	↑	ATM(↑), URB(↑)	0.19	0.34	0.55	1	Reject
% Planktonic	0.00	66.76	11.47	log(x + 1)	↑		0.01	0.13	0.10		Reject
% Eutrophic	0.00	71.14	8.71	log(x + 1)	↑	AG(↑), IND(↑), URB(↑)	0.21	0.57	0.36	3	Reject
% Chl *a* tolerant	0.00	79.84	2.90	log(x + 1)	↑	AG(↑)	0.18	0.56	0.33	3	Reject
% TSS tolerant	0.00	75.33	8.02	log(x + 1)	↑	URB(↑)	0.16	0.49	0.32	3	Reject
% 1/TTube tolerant	0.00	72.94	2.16	log(x + 1)	↑	ATM(↑), URB(↑)	0.20	0.52	0.38	3	Reject
% Cl tolerant	0.00	71.09	8.63	log(x + 1)	↑	AG(↑), ATM(↑), IND(↑), URB(↑)	0.35	0.65	0.54	3	Reject
% Araphid	0.00	88.26	21.23	log(x + 1)	?	ATM(↓)	0.19	0.37	0.50	2	Accept
% Monoraphid	0.00	81.19	29.34	None	?	AG(↓)	0.10	0.26	0.39		Accept
% Biraphid	5.03	90.24	43.29	None	?	ATM(↑)	0.17	0.29	0.60	1	Accept
% N fixers	0.00	5.57	0.18	log(x + 1)	↓		0.01	0.07	0.13		Reject
% N heterotrophs	0.00	38.21	8.48	log(x + 1)	↑		0.04	0.20	0.22		Reject
% Stephanodiscoid	0.00	27.32	1.48	log(x + 1)	↑		0.13	0.36	0.36	3	Reject
% *Cyclotella*	0.00	17.30	1.40	log(x + 1)	↓		0.02	0.17	0.13		Reject
% *Staurosira* complex	0.00	87.78	14.24	log(x + 1)	?	AG(↑), ATM(↓), URB(↓)	0.17	0.38	0.46	2	Reject
% *Fragilaria* + *Synedra*	0.00	60.74	5.02	log(x + 1)	?		0.03	0.06	0.50		Reject
% *Martyana*	0.00	18.40	0.95	log(x + 1)	?	ATM(↓)	0.11	0.27	0.41		Accept
% *Eunotia*	0.00	36.89	0.74	log(x + 1)	?	AG(↓)	0.06	0.26	0.24		Accept
% *Achnanthidium minutissimum* complex	0.00	72.62	10.08	log(x + 1)	↓	IND(↓)	0.07	0.22	0.30	5	Accept
% *Cocconeis*	0.00	57.25	5.28	log(x + 1)	?	ATM(↑)	0.09	0.10	0.89		Accept
% *Planothidium*	0.00	51.29	7.51	log(x + 1)	?		0.02	0.12	0.15	5	Reject
% *Psammothidium*	0.00	35.31	1.66	log(x + 1)	?	AG(↓)	0.22	0.30	0.72		Accept
% *Rossithidium*	0.00	10.72	1.10	log(x + 1)	?	URB(↓)	0.05	0.11	0.39		Accept
% *Karayevia*	0.00	8.58	0.67	log(x + 1)	?		0.02	0.12	0.18		Reject
% *Navicula*	0.24	53.12	10.17	log(x + 1)	?	ATM(↑)	0.21	0.34	0.62		Accept
% *Geissleria*	0.00	36.78	2.59	log(x + 1)	?	AG(↓), ATM(↑), URB(↑)	0.09	0.28	0.30		Accept
% *Hippodonta*	0.00	39.53	1.38	log(x + 1)	?	ATM(↑), IND(↑), URB(↑)	0.18	0.27	0.67		Accept
% *Sellaphora*	0.00	32.92	3.43	log(x + 1)	?		0.00	0.19	0.02		Reject

TABLE 1. (Continued)

Metric	Minimum	Maximum	Average	Transformation applied	Anticipated response to disturbance	Actual response to disturbance	$R^2_{\text{anthropogenic}}$	R^2_{natural}	$R^2_{\text{anthropogenic}}$:$R^2_{\text{natural}}$ ratio	Cluster	Decision
% Cymbelloid	0.00	35.71	3.90	$\log(x + 1)$	?		0.04	0.23	0.17	5	Reject
% *Gomphonema*	0.00	60.41	2.74	$\log(x + 1)$	?		0.02	0.15	0.13		Reject
% *Amphora*	0.00	41.67	7.20	$\log(x + 1)$	?	AG(↑)	0.06	0.18	0.32		Accept
% *Nitzschia*	0.00	41.69	5.23	$\log(x + 1)$	?		0.05	0.33	0.15		Reject

Basic statistics are provided for each metric, and also shown are the data transformations applied to meet statistical assumptions of the multiple regression procedure. Where possible, anticipated responses to anthropogenic disturbance were determined from previous literature. Actual responses were determined from the multiple linear regression coefficients against the stressor variables agriculture (AG), atmospheric deposition (ATM), urban development (URB), and industrial point sources (IND). Anticipated and actual responses to increased disturbance are illustrated as increasing (↑), decreasing (↓), or unknown (?). Squared correlation coefficients reflect multiple linear regression relationships between candidate metrics and anthropogenic variables ($R^2_{\text{anthropogenic}}$) or natural variables ($R^2_{\text{natural}}$). Cluster numbers indicate groups of highly correlated variables, as illustrated in Figure 2. The decision column indicates whether a candidate metric was accepted (based on selection criteria) for inclusion in the final multimetric index.

DCA, detrended correspondence analysis; DI, diatom-inferred; TP, total phosphorus; TSS, total suspended solids; 1/TTube, inverted transparency tube measurement.

dominance by a single taxon would increase with greater environmental disturbance.

Proportion motile diatoms: The proportion of motile diatoms has been used as an index of siltation in Montana streams (Bahls 1993) and cited as a potential metric by other investigators (Mills et al. 1993, Stevenson and Pan 1999, Hill et al. 2000, Fore and Grafe 2002). Periphytic diatoms with the ability to move purposely using their raphe structure might have a competitive ability in systems where anthropogenic activities periodically generate silt and other factors that influence water clarity and/or sedimentation. We expected that the motile diatom metric would increase with greater environmental disturbance.

Proportion planktonic diatoms: Studies of contemporary and sediment core assemblages from lakes have shown that planktonic diatoms increase in relative abundance in response to cultural eutrophication (Nygaard 1949). Other diatom studies have further shown that there are characteristic oligotrophic (e.g., *Cyclotella*; Stockner and Armstrong 1971) and eutrophic (e.g., *Cyclostephanos*; Reavie et al. 2006) species of planktonic diatoms, subgroups of the plankton that have opposite meaning in terms of nutrient enrichment. We anticipated that planktonic diatoms would increase in relative abundance with increased disturbance, particularly at disturbed sites in the productive Lake Erie.

Proportions of "tolerant" diatoms: A number of published studies provide index values for diatom species that reflect their tolerance and sensitivity to water quality variables, such as nutrients (e.g., Van Dam et al. 1994, developed in the Netherlands). For example, Fore and Grafe (2002) used published lists of autecological characteristics of diatom species to derive metrics for Idaho rivers. Because detailed water chemistry data were available for our Great Lakes coastal locations, we derived new lists that describe the species responses to water quality variables. Similar to those selected for DI data (above), TP, chl *a*, Cl, TSS, and 1/TTube were identified as strong determinants of diatom assemblage properties (Reavie et al. 2006) and potentially important indicators of anthropogenic influence. "Tolerant" and "sensitive" classifications were developed for the common diatom species for each of these five variables (Appendix S1 in the supplementary material), and metrics were calculated based on the proportion of tolerant species in each sample assemblage.

To characterize the autecological information for each taxon, species optima and tolerances were derived as described above for transfer function development (Reavie et al. 2006). Species with autecological characteristics greater than selected cut-off criteria were identified as being tolerant to that environmental variable. We started with TP, setting a cut-off criterion of 30 $\mu g \cdot L^{-1}$, a standard definition for nutrient concentrations in eutrophic systems and used by many states for assessment and regulatory purposes (Carlson 1977, MDEQ 2001, MPCA 2004a,b, WDNR 2005). We took a conservative approach to avoid the inclusion of cosmopolitan species (i.e., species with broad tolerances covering a large portion of the environmental gradient); species with a lower tolerance limit greater than the cut-off value were identified as tolerant. Using this criterion, 20% of the species were identified as tolerant of high TP concentrations. Similar methods were used to identify species that are tolerant of high turbidity (1/TTube), chl *a*, TSS, and Cl concentrations, but tolerant species were considered those that had lower tolerance limits above the 80th percentile of species (i.e., the 20% most tolerant taxa). In this way, the respective metrics would similarly define the upper 20% most tolerant species in the Great Lakes coastal systems. The critical concentrations at the 80th percentile were 16.0 $\mu g \cdot L^{-1}$ for chl *a*, 5.4 $mg \cdot L^{-1}$ for TSS, 0.010 cm^{-1} (98 cm tube visibility) for 1/TTube, and 11.5 $mg \cdot L^{-1}$ for Cl. For a user, these five metrics are relatively easy to calculate on the basis of species assignments in Appendix S1.

Proportions of araphid, monoraphid, and biraphid diatoms: The pennate diatoms can be divided into three distinct groups based on the presence of the raphe, a longitudinal slit through the valve face that is characteristic to several pennate genera. The raphe is used for motility, adherence, and in some species appears to be superfluous (Round and Crawford 1990). Araphid genera (e.g., *Staurosira*, *Fragilaria*) have no raphe structure and tend to dominate the pennate-planktonic and epipsammic taxa. Monoraphid genera (e.g., *Achnanthidium*, *Cocconeis*) have a raphe on one valve and are common epiphytic taxa. Biraphid genera (e.g., *Navicula*, *Nitzschia*) have a raphe on both valves and are common in all periphytic environments. To date, such metrics have not been rigorously tested, and we were uncertain about relationships between these diatom groups and stressor data.

Proportions of nitrogen fixers and heterotrophs: The diatom genera *Epithemia* and *Rhopalodia* are known nitrogen fixers because they harbor endosymbiotic bacteria that can convert atmospheric nitrogen (N_2) into biologically accessible forms (Mulholland 1996). Diatom species such as *Nitzschia frustulum* and *Mayameae atomus* (see Appendix S1 for taxonomic authors) are known heterotrophs and so under certain conditions can use amino acids and other metabolites created by other organisms as a source of nutrients (Tuchman 1996). The proportions of nitrogen fixers and obligate + facultative nitrogen-heterotrophic taxa were calculated based on species assignments developed by Van Dam et al. (1994). Although little is known about the ecological significance of these characteristics of certain diatoms, we anticipated that nitrogen fixers would decline and nitrogen heterotrophs would increase with increasing stress that would increase both organic and inorganic nitrogen loading to the water column.

Proportions of species and species groups: Several metrics were derived from 15 sufficiently abundant genera to determine if they had potential as metrics. We did not anticipate the direction of responses of periphytic genera such as *Cocconeis* and *Planothidium* to anthropogenic stressors, so their evaluation was treated as exploratory (details of expected responses, based on the literature, are in Table 1). Metrics based on groups of genera were also defined: stephanodiscoid (*Stephanodiscus* + *Cyclostephanos*), *Staurosira*-type complex (*Staurosira* + *Staurosirella* + *Pseudostaurosira*), and cymbelloid (*Cymbella* + *Encyonema* + *Encyonopsis* + *Cymbopleura* + *Delicata* + *Cymbellopsis* + *Afrocymbella*). *Stephanodiscus* and *Cyclostephanos* are generally considered planktonic indicators of high nutrient concentrations (Anderson 1990), so it was expected that the relative abundance of stephanodiscoid diatoms would increase with increasing stress.

Cyclotella has often been attributed to oligotrophic open-water conditions (Stockner and Armstrong 1971), but certain species of *Cyclotella*, such as *C. stelligeroides*, have been noted to be indicators of high nutrients. Our description of *Cyclotella* includes some likely polyphyletic genera that have been differentiated, including *Discostella* and *Puncticulata*. We maintained *Cyclotella* as a simplified index but acknowledge that newer information on the polyphyletic nature of this genus may indicate that it contains several genera with different autecological properties.

We could not speculate on the responses of araphid and cymbelloid complexes to stress, so they were treated as exploratory metrics. Because it was the most common taxon identified across the Great Lakes samples, a metric ''*Achnanthidium minutissimum* complex'' was derived based on the sum of the relative abundance of *A. minutissimum*, its forms and varieties, as well as taxa formerly belonging to the species definition of *Achnanthes minutissima* (Krammer and Lange-Bertalot 1991). We chose to build this complex because the subspecies identification for *A. minutissimum* is often not possible using LM, particularly when specimens are enumerated in girdle view. Results from the Great Lakes indicate a clear dominance of *A. minutissimum* in lower-nutrient environments (Reavie et al. 2006), so we anticipated that this metric would decrease with greater environmental disturbance.

Quantifying upland anthropogenic and natural watershed characteristics. The following methods briefly summarize approaches used by Danz et al. (2005), Reavie et al. (2006), and T. Hollenhorst (personal communication) to characterize sample sites. Within GIS, a polygon was drawn encompassing sampling points for all GLEI indicator groups at a selected locale, and this polygon was assumed to be the receiving area for the watershed. Watersheds were delineated for each of the wetland and embayment polygons using 30 m digital elevation models and ArcInfo (ESRI 2000). High-energy sample locales have some fundamental physical differences from the wetlands and embayments, particularly that the biota are more exposed to physical activity in the adjacent Great Lake, and those biota are likely to be influenced by a longer stretch of receiving coastline than that defined by the polygons. Therefore, identifying the actual contributing area for high-energy sites required some additional calculations (T. Hollenhorst, personal communication). First, the immediate watershed for each high-energy polygon was agglomerated with watersheds on either side, including stream catchments and coastal catchments with no streams (''interfluves''). Agglomeration continued on each side of the polygon's watershed until the threshold area of 9 km^2 was reached; this threshold was determined as twice the area of the median segment-shed size across the lakes. The associated stressor summaries were then summarized for each agglomerated high-energy watershed area using area-weighted means for the corresponding watershed areas.

Over 200 environmental variables in seven categories of environmental variation (Danz et al. 2007) were summarized for each watershed. Principal components analysis (PCA) within five categories of environmental variation was used to reduce dimensionality and derive overall gradients. For example, 26 agricultural variables (including pesticide runoff and leaching, cropland area, nitrogen and phosphorous exports, percent of county treated for various pests, and livestock inventories) comprised an agricultural category, and the first principal component (PC) was identified as a gradient of watershed agricultural activity for the Great Lakes coastlines. The final watershed-level predictor matrix consisted of four anthropogenic principal components (agriculture [AG], atmospheric deposition [ATM], industrial point source pollution [IND], and urbanization [URB]) and one natural PC (soil characteristics [SOIL]).

Identifying metrics from candidate metrics. Several candidate metrics had skewed distributions, and $\log_{10}$ transformation was used as appropriate to better meet statistical assumptions (Table 1). The following approach was used to identify metrics for multimetric development from the complete list of candidate metrics: (i) Correlation among candidate metrics was investigated to identify and eliminate redundancy. (ii) The suitability of each candidate metric was evaluated using stepwise regression to stressor principal components. (iii) Similarly, each candidate metric was related to natural gradients using stepwise regression to identify metrics that were being largely determined by natural factors. Within groups of covarying candidate metrics, metrics were selected that best tracked stress and least tracked natural gradients.

Correlation among candidate metrics: All candidate metrics were compared using a pair-wise correlation matrix. Candidate metric pairs with an absolute-value Pearson correlation coefficient ($|r|$) greater than 0.7 were noted as being redundant. Candidate metrics were also explored using PCA to provide a comprehensive assessment of the relationships among these metrics.

Candidate metric/stressor relationships: Each candidate metric was regressed against stressor data using multiple linear regression and evaluated using the squared Pearson's correlation coefficient (R^2). Four stressor variables were used in multiple regressions: agricultural (AG), atmospheric (ATM), industrial (IND), and urban (URB) PC axis scores. This regression tested the relationship between watershed properties and diatom metrics in the adjacent coastal system and was used to determine which candidate metrics were being significantly influenced by watershed stressors. A metric with a significant relationship to stressors would have potential as a component of the final MMDI.

Preliminary visual assessment of metric-stressor scatterplots, as recommended by Karr and Chu (1999), indicated our assumption of linear relationships between metric and stressor data was appropriate in most cases; a majority of these plots showed either linear relationships or no noticeable pattern. A few of the metrics had subtle threshold (i.e., varying in response above and below a particular stressor value) or wedge-shaped (i.e., narrowing or widening of the range of metric scores as stressor values increased) responses along stressor gradients. Despite the presence of some complicated metric responses, we decided that assuming linear responses was still the most appropriate, user-friendly choice for development of the index.

Candidate metric/natural gradient relationships: Similar to characterizing relationships to stressors, candidate metrics were related to a suite of variables describing the natural variation that occurred in our data set because of spatial, geophysical, and time factors. These natural variables were latitude, longitude, watershed soil PC scores, habitat type, and water temperature. For regression of habitat, five binary (dummy) variables were created, one for each coastal habitat type. For instance, a coastal wetland (CW) variable was created; samples from coastal wetlands were identified by ones, and all other habitats were given zero values. Water temperature was assumed to be a seasonal variable that also may have been influenced by geomorphology. For instance, in midsummer, a shallow, protected wetland is likely to be warmer than a high-energy locale that is more directly influenced by water from the adjacent Great Lake.

Selection of final metrics: Metrics were selected if they did not strongly covary with other candidate metrics and were significantly correlated to one or more watershed stressor variables. Within clusters of candidate metrics (identified by the correlation matrix), the metric with the highest ratio of stressor/natural gradient correlation was selected, assuming a significant relationship to the stressor gradient existed for that variable.

Multimetric development. The MMDI was developed based on the sum of the selected metrics, with each metric weighted based on the strength of its relationship to stressors. In this way, metrics with weak, but still significant, relationships to anthropogenic stressors would play a lesser role in multimetric calculations.

Part of the GLEI project involved the identification of reference locations for this and future ecological assessments (Host et al. 2005). Using remotely sensed and other GIS data, the degree of anthropogenic disturbance was characterized for a suite of discrete polygons spanning the U.S. shoreline of the Great Lakes. These polygons were assigned a maximum relative score (MAXREL) based on their contributing stressors. Although the goal of Host et al. (2005) was to identify the ''least disturbed'' (i.e., lowest MAXREL) sites for reference considerations, all sites were assigned a ''reference score.'' Recent refinement of this index involved standardization of the composite stressors and summing of these stressors into a new score, SUMREL (G. E. Host, L. B. Johnson, J. J. H. Ciborowski, T. P. Hollenhorst, unpublished; G. Host, Natural Resources Research Institute, personal communication). Because these SUMREL sites encompass our diatom sample locations, we were able to regress reference scores against selected metrics and MMDI scores, providing us with a means to assess the ability of the MMDI to track stressor influences.

To test the ability of the MMDI and its component metrics to track disturbance, a test set of 47 samples (30%) was selected, and the index (including all component metrics) was redeveloped using the new 108-sample data set. Samples in the test set were selected to maximize the gradient of anthropogenic influence; all samples were ordered according to their agricultural principal component score, and every third sample was picked for testing. Using the reduced ''calibration'' data set, new metric and MMDI equations were developed, and scores were calculated for each of the test samples. Test sample scores were related to corresponding SUMREL scores to quantify the relative strength of the metrics and MMDI to reflect ecosystem stressors.

RESULTS

Identifying metrics from candidate metrics. Correlation among candidate metrics: A correlation matrix (Appendix S2 in the supplementary material) contained numerous significant correlations among the candidate metrics, but relatively few correlations above $|r| = 0.70$. Not surprisingly, particularly high correlation is identified between DI chl *a* and DI TP, two variables that strongly reflect trophic condition. Other correlated pairs that were expected include the Shannon–Weaver calculations based on all taxa and genera, and % araphid/% *Staurosira* complex; *Staurosira* and its related genera are araphid and in many cases make up the bulk of the araphid diatom assemblages in the coastal Great Lakes samples.

The PCA of candidate metrics illustrates the clusters of intercorrelated metrics (Fig. 2); axes 1 and 2 account for 36% and 21% of the variance, respectively. Five clusters of intercorrelated candidate metrics were identified, and 19 metrics were not considered highly redundant with any others (Table 1). Percent motile taxa was grouped with percent biraphid taxa (cluster 1), an obvious result of the fact that the majority of freshwater biraphid taxa are also motile. Interestingly, percent *Navicula* (a motile genus) was not considered redundant with these two metrics, although it occurs nearby with a relatively high axis 1 score. On the opposite quadrant from the motile-biraphid group lies nonmotile metrics, araphid taxa and the *Staurosira* complex (cluster 2). The upper right quadrant contains a large cluster of intercorrelated candidate metrics that are known to reflect disturbance (cluster 3). Although DCA AX1 is not intuitively related to the other metrics in the cluster, it is not surprising that the primary axis in the DCA is largely a gradient of low to high trophic condition or disturbance and so is correlated with other disturbance metrics. Cluster 4 represents three diversity-based metrics. The first two axes of the PCA (Fig. 2) do not adequately show the separation of this cluster from the other

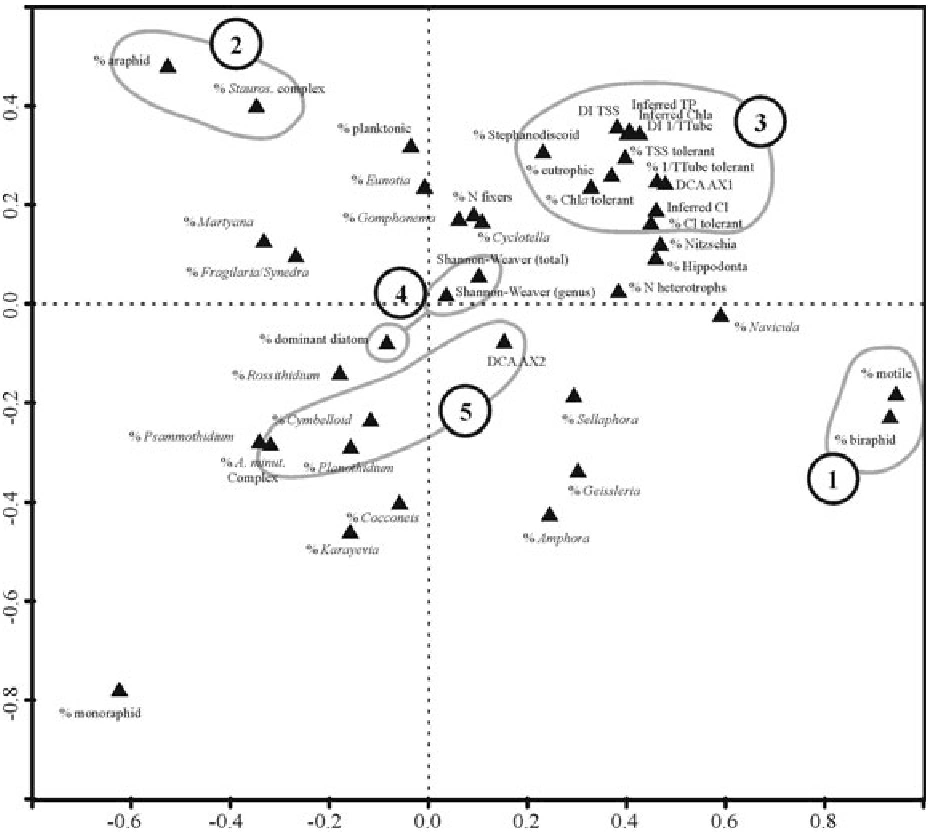

FIG. 2. Principal components analysis (PCA) of the candidate metrics. Groups of redundant metrics, as determined using a correlation matrix, are encircled. The cluster containing anticorrelated variables is grouped via a connecting line through the origin. 1/TTube, inverted transparency tube measurement; Cl, chloride; DCA, detrended correspondence analysis; DI, diatom-inferred; TP, total phosphorus; TSS, total suspended solids; *A. minut.*, *Achnanthidium minutissimum*; *Stauros.*, *Staurosira*.

metrics near the center of the diagram. The members of this ''diversity'' cluster all have high PCA axis 3 scores, and so this cluster can be considered to be separated from these other metrics along the third axis, which is not shown. The members of cluster 5 do not have an intuitive relationship but may reflect occurrences of ecologically similar diatom taxa.

Regressions against stressors: Several interesting metric responses to stressors were observed (Table 1). Multiple regression identified significant relationships between 26 candidate metrics and the four stressor variables. Several of these responses were expected, such as the increase in relative abundance of eutrophic, TSS-tolerant, and Cl-tolerant taxa with increasing stress. Strong relationships were found between candidate metrics and each of AG, ATM, URB, and IND. The strongest overall response was increasing DI Cl against increasing agricultural intensity. It can also be seen that there were significant responses of metrics such as % biraphid diatoms and % *Martyana* to stressors including urban development and atmospheric deposition, respectively.

Regressions against natural variables: Based on multiple linear regressions with the natural Great Lakes gradients (R^2_{natural}; Table 1), it is clear that the candidate metric scores are more strongly driven by the natural factors we measured than the four stressor variables. The ratio of $R^2_{\text{anthropogenic}}$ to R^2_{natural} indicates that the relative response to stressors ranged from 0.10 (% planktonic) to 0.89 (% *Cocconeis*), although it is noteworthy that candidate metrics with high $R^2_{\text{anthropogenic}}$ values also tended to have high R^2_{natural} values.

Selection of metrics from candidate metrics: Fourteen of the candidate metrics were considered both responsive to stressors and nonredundant and so met the selection criteria for multimetric development (Table 2). There was moderate variation in the relative weights of selected metrics. For instance, the weaker response of the % *Martyana* to stressors (Table 1) means % *Martyana* will be given a lesser weight than, for example, DI Cl in the MMDI.

MMDI development. One of the selected metrics, % *Geissleria*, had conflicting responses to stressor variables (Table 1). In this case, the stress-response to be used in multimetric calculation was chosen to be positive (i.e., increasing in value with greater stress) based on independent *t*-tests during multiple regression calculations ($t = -2.01$ for AG, 2.16 for ATM, and 3.02 for URB). Because of the dominant

TABLE 2. Metrics selected for inclusion in the multimetric index, including weights and scoring equations to be used in multimetric calculations.

Metric	Transformation applied	Response to stress	Weight (w)	Score (M) calculation
DI Cl (log [mg · L^{-1} + 1])*	None	↑	2.311	$1.094 - x/1.71$
% Araphid	Log	↓	0.915	$\log(x+1)/1.95$
% Monoraphid	None	↓	0.491	$x/81.19$
% Biraphid	None	↑	0.832	$1.059 - x/85.21$
% *Martyana*	Log	↓	0.550	$\log(x+1)/1.29$
% *Eunotia*	Log	↓	0.302	$\log(x+1)/1.58$
% *Achnanthidium minutissimum* complex	Log	↓	0.326	$\log(x+1)/1.87$
% *Cocconeis*	Log	↑	0.414	$1 - \log(x+1)/1.77$
% *Psammothidium*	Log	↓	1.061	$\log(x+1)/1.56$
% *Rossithidium*	Log	↓	0.219	$\log(x+1)/1.07$
% *Navicula*	Log	↑	1.007	$1 - [\log(x+1) - 0.093)]/1.64$
% *Geissleria*	Log	↑	0.418	$1 - \log(x+1)/1.58$
% *Hippodonta*	Log	↑	0.876	$1 - \log(x+1)/1.61$
% *Amphora*	Log	↑	0.277	$1 - \log(x+1)/1.63$

Diatom-inferred (DI) chloride (Cl) (*) requires high-resolution taxonomy and so may be removed in a simplified application of the multimetric application.

positive responses of % *Geissleria* to the ATM and URB stressor variables, this response was the chosen assumption in multimetric calculations. It was necessary to standardize the selected metric scores using the following equations as appropriate. A raw score for a metric was calculated as a value between zero and one, with higher values reflecting less stress (i.e., better watershed conditions). If a metric responded positively to stress, the metric's score (M) was calculated using the following equation:

$$M_n = 1 - \left(\frac{x_n - \min_n}{\max_n - \min_n}\right) \quad (1)$$

where n is the metric identifier, $\min_n$ and $\max_n$ are the minimum and maximum measured values for that metric from the Great Lakes training set, and x_n is the measured metric value at a given site. The score was subtracted from one to invert the value, such that previously higher values become lower to reflect ''poorer'' conditions. Metrics with negative responses to stress do not require this correction and so were calculated using equation 2.

$$M_n = \left(\frac{x_n - \min_n}{\max_n - \min_n}\right) \quad (2)$$

The resultant scores for each selected metric were then multiplied by their respective weight (w) and summed to create the multimetric total, G (eq. 3). This equation assumes the 14 selected metrics (Table 2) are being used.

$$G = \sum_{n=1}^{14} (M_n) \times (w_n) \quad (3)$$

All metric calculations were scaled to create an index that, theoretically, lies between 0 and 10, with 0 being the least desirable condition and 10 being the best. Scaling was performed by summing (total of selected $R^2_{\text{anthropogenic}}$ = 2.06) and scaling up the weights (w) so that this total became 10, the theoretical highest G. Of course, because the highest Ms did not all occur at the same coastal site, the maximum G from the training set was not expected to reach 10. However, G values outside the range of those calculated from the training set may be obtained in the future, as metric scores may be less or greater than those obtained in our training set.

Final calculation requirements for each M are shown in Table 2. For example, the calculation for % *Navicula* ($x_{\%\ \text{Navicula}}$) is calculated using the minimum (0.24%) and maximum (53.1%). Also note that the ''1-'' correction is needed because of the positive response of % *Navicula* to stress, and that it is required that the input values be log-transformed for this metric (1 was consistently added to log-transformed data due to the presence of values <1).

$$M_{\%\text{Navicula}} = 1 - \left(\frac{\log(x_{\%\text{Navicula}} + 1) - \log(0.24\% + 1)}{\log(53.1\% + 1) - \log(0.24\% + 1)}\right) \quad (4)$$

Reducing equation 4 results in the equation as shown in Table 2. Weights are also provided, and note that the weight for % *Navicula* has been scaled from 0.207 ($R^2_{\text{anthropogenic}}$) up to 1.007. As denoted by equation 3, the series of weight-times-score calculations were then summed for a given sample, providing the final multimetric score.

A significant negative correlation occurred between the MMDI and SUMREL reference scores (Fig. 3A). Clearly, a higher MMDI score equates with a lower SUMREL (i.e., lower stress), indicating that the MMDI provides appropriately low results in reference areas. However, when the MMDI was applied to the 43-sample test data set, there was a decline in the squared correlation coefficient between MMDI and SUMREL, from $r^2 = 0.54$ to $r^2 = 0.32$. Three of the candidate metrics with high $r^2_{\text{anthropogenic}}$ (Table 1) were also selected, and these

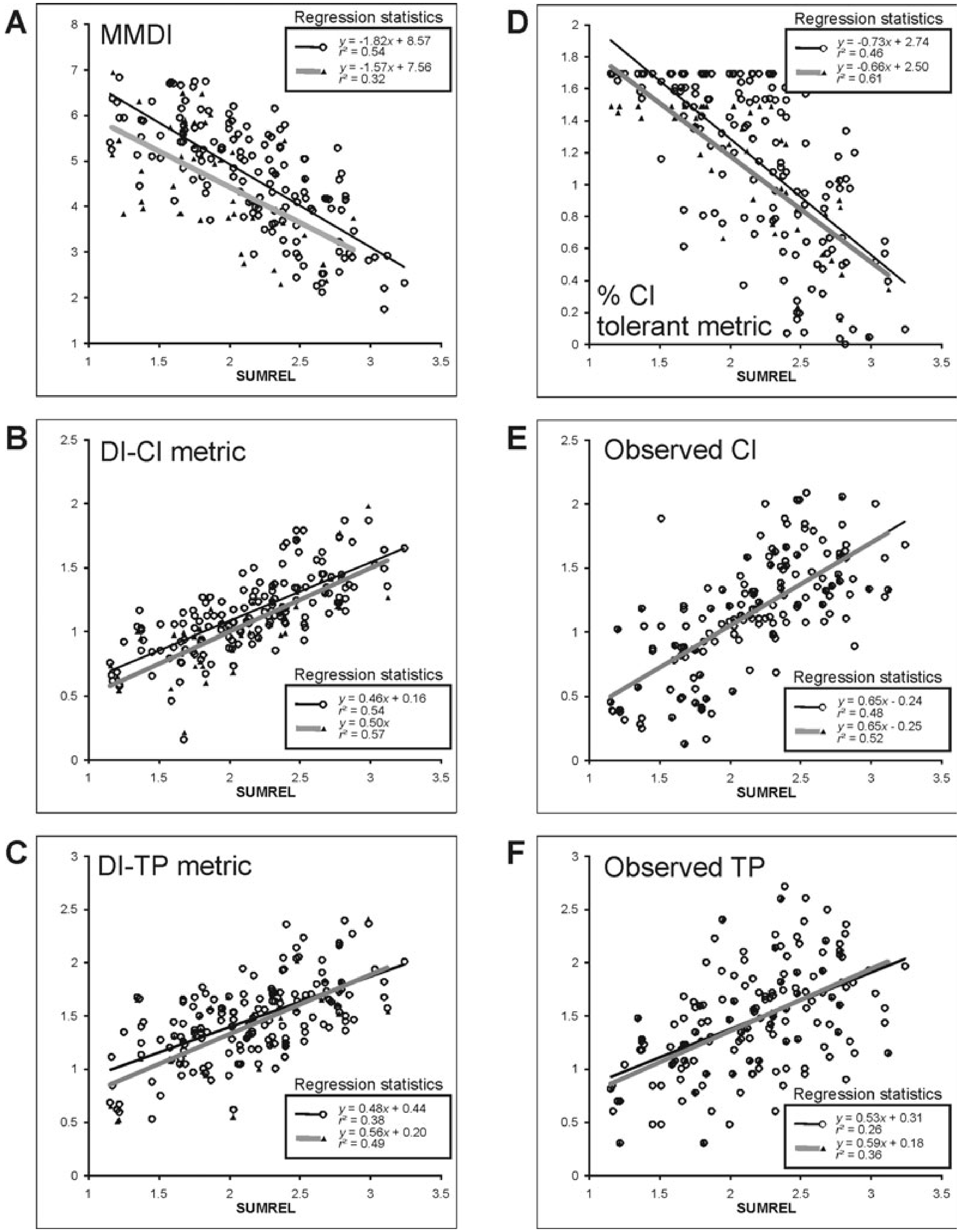

FIG. 3. Regressions of MMDI and selected metric scores against SUMREL (an indicator of watershed stressor influence). Regressions are illustrated for the full data set (155 sites, black lines) and independent test data set (47 sites, gray lines). Cl, chloride; DI, diatom-inferred; MMDI, multimetric diatom index; SUMREL, standardized composite stressor value; TP, total phosphorus.

candidate metrics were also strong predictors of SUMREL. There was no decline in the ability of DI-Cl (r^2[full sample set] = 0.54, r^2[test sample set] = 0.57; Fig. 3B), DI-TP (r^2[full sample set] = 0.38, r^2[test sample set] = 0.49; Fig. 3C) or % Cl-tolerant (r^2[full sample set] = 0.46, r^2[test sample set] = 0.61; Fig. 3D) to track stressors at the test locations. In no comparisons were the slopes of full and test regressions significantly different (Student's *t*-test, $P = 0.05$).

To determine if using diatom-based reconstructions provided better information about watershed stressors than measured chemical variables, SUMREL was also compared with measured data. Both DI-Cl (r^2[test sample set] = 0.57; Fig. 3B) and DI-TP (r^2[test sample set] = 0.49; Fig. 3C) were better correlated to SUMREL than measured Cl (r^2[test sample set] = 0.52; Fig. 3E) and TP (r^2[test sample set] = 0.36; Fig. 3F), respectively. It is interesting that for the three metrics we tested (Fig. 3, B–D), r^2 consistently increased in the test sample set regressions. The same thing occurred for observed Cl and TP (Fig. 3, E and F), indicating that at least part of this increase was likely due to the specific subset of sample data chosen for independent testing.

Metric score distribution across the Great Lakes. There is a clear distinction in the metric scores between the southern (EBF) and northern (LMF) ecoprovinces (Fig. 4).

DISCUSSION

This manuscript provides a method for metric and multimetric development using algae and a set of detailed, previously unavailable stressor data for the Great Lakes. Although widely used, multimetric approaches have been scrutinized because of the subjectivity involved in metric selection, questionable metric responses to stressors, the implications of covariance among metrics, and the validity of compiling metric scores as if each metric was equally important (Fore et al. 1996). Also, index approaches leading to enforceable regulations must be able to distinguish between human-caused changes in the biological assemblage and natural variability. In this study, we made efforts to take these concerns into account while analyzing diatom responses to anthropogenic variables. We have attempted to minimize many of the problems that have been associated with multimetric approaches: we assessed metric responses to robustly defined anthropogenic and natural gradients, and we down-weighted or eliminated metrics that were less able to track stressors. It was necessary to eliminate some metrics that are traditionally associated with anthropogenic impacts that are commonly applied in biological monitoring criteria. For instance, species diversity is an archetypal assessment of habitat quality for many organism groups (Huston 1979), but for the Great Lakes diatoms, the Shannon–Weaver metric appears to be of little use. Functional (e.g., % araphid) and genera (e.g., % *Navicula*) groups were much more meaningful in terms of tracking stressors and so comprised much of the multimetric tool.

We hypothesized that combining numerous complementary indicator metrics would provide a robust means to infer condition. While the MMDI provided scores that were related to watershed stressors, it did not provide an advantage over the best of the candidate metrics. Even with down-weighting of the weaker metrics, these metrics impaired the collective ability of the MMDI. At this time, we do not recommend this diatom-based multimetric approach for Great Lakes coastlines. Even a relatively simple metric such as % Cl tolerant taxa reflected watershed stressors better than the MMDI. Metrics based on weighted averaging (DI-Cl, DI-TP; Reavie et al. 2006) remain the superlatives because of their proven ability to reflect stress and their superiority over their measured counterparts (e.g., water quality measurements for Cl and TP).

How should these metrics be used? Several useful metrics have resulted from this work, and Figure 4 includes 12 of the metrics that we feel have some value as indicators in monitoring programs. Although there is much intercorrelation among metrics, a user may choose certain metrics to reflect particular aspects of site quality. For instance, % Cl-tolerant taxa infers relative impacts from road salt applications, whereas % eutrophic taxa infers nutrient characteristics. Other metrics, such as % *Navicula* and DCA axis 1 score, are less specific and would be used as more general indicators of stress.

The following are recommendations for interpretation of metric results, acknowledging that cut-off criteria for what might be considered ''good'' or ''poor'' condition is rather arbitrary. Because of the differences in geology, population density, and types of human activities between the LMF and EBF, results from the two ecoprovinces should be treated separately, using percentile data (e.g., Fig. 4) as a general guide. Using % Cl-tolerant species as an example, and in the context of our data set, scores greater than 1.4 (~24% Cl-tolerant taxa, back-transformed) from EBF represent the top 25th percentile (i.e., the greatest stressor impact). This upper 25th percentile in the LMF includes scores >~0.4 (~1.5%). The least impacted sites would be those <0.5 in the EBF, or zero in the LMF.

We anticipate application and possible refinement of these diatom-based indices in the future. We expect that the value of these metrics will be illustrated from their eventual use as tools in paleoecological or monitoring applications. Although we were unable to develop a sufficiently robust multimetric index, we do not recommend against such an approach in the future. It may be that a more

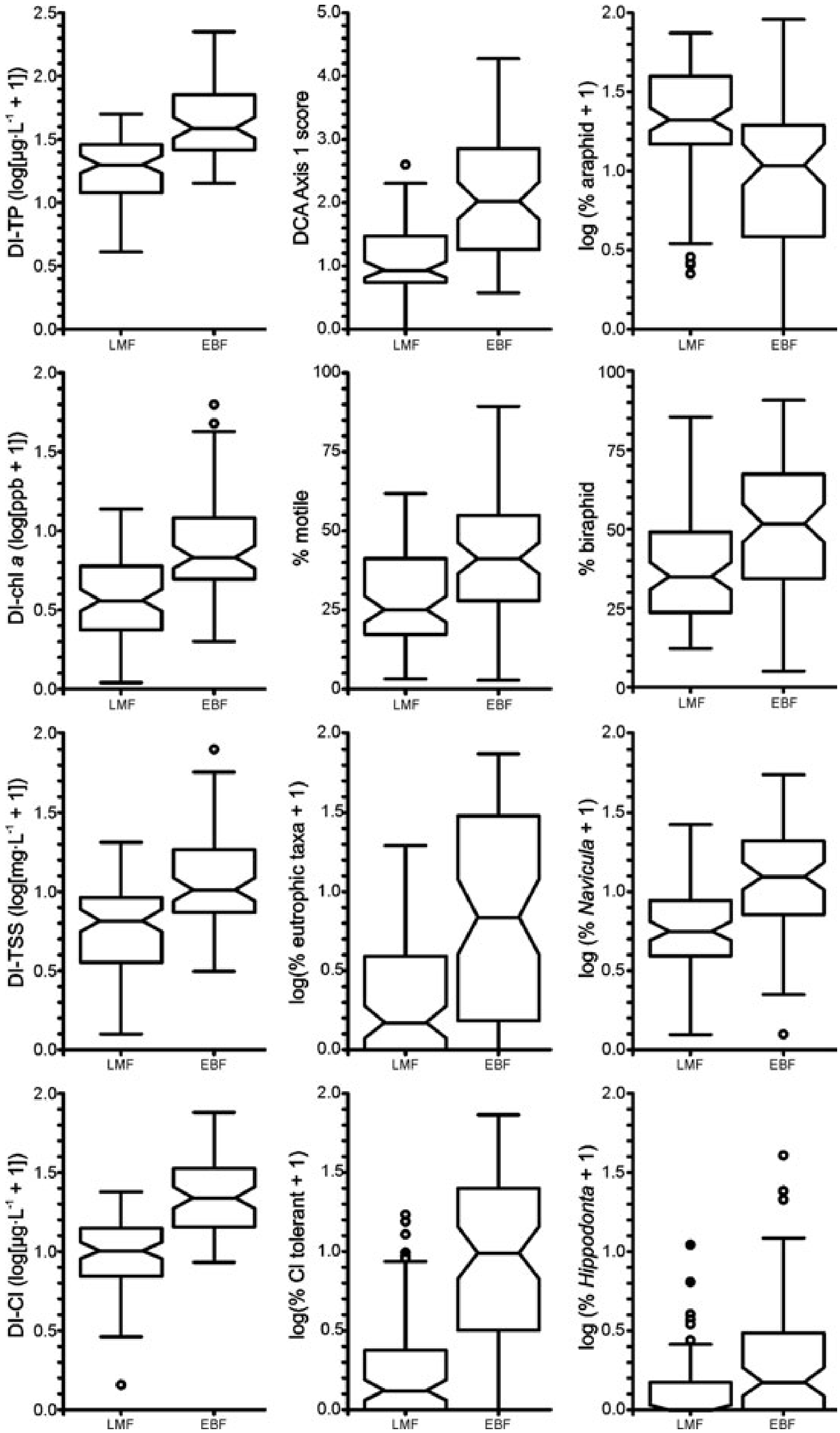

FIG. 4. Boxplots of selected metric scores from the Great Lakes coastal training set diatom assemblages. Metrics were selected for this plot of they had $R^2_{\text{anthropogenic}} > 0.15$ and $R^2_{\text{anthropogenic}}:R^2_{\text{natural}}$ ratio > 0.35, indicating that they had relatively strong relationships with anthropogenic factors. Plots illustrate score ranges for samples in the northern (Laurentian mixed forest; LMF) and southern (eastern broadleaf forest; EBF) ecoprovinces. The top and bottom of each box are the 75th and 25th percentiles, and a line is drawn through the middle of each box at the median. Notches in each box are used for statistical comparison; if notches from two boxes do not overlap, the medians are significantly different ($P < 0.05$). The upper and lower tails, respectively, indicate the largest and smallest scores adjacent to the 1.5 interquartile ranges of the upper and lower box percentiles. Circles indicate outliers. Cl, chloride; DCA, detrended correspondence analysis; DI, diatom-inferred; TSS, total suspended solids.

stringent selection process is needed for component metrics, and/or that the calculated outputs from individual metrics need to be considered for their independent ecological meaning, instead of lumping them into a single index.

This work establishes new (and tests existing) methods for algal indicator development and valuable tools for interpretation of monitoring data and paleolimnological records. We used independent assessments in two ecoprovinces to ensure that our selected metrics were at least able to distinguish a broad gradient of anthropogenic disturbance, and we look forward to continued index application, validation, and refinement in the future.

This work was made possible through the efforts of many GLEI and EPA researchers who assisted in the field, laboratory, and with project logistics. We are grateful to V. Brady for coordinating and assisting with the sampling design. Diatom identification and enumeration were supported by N. Andresen and M. Ferguson, and diatom taxonomic support was provided by E. Stoermer and J. Johansen; J. Henneck was a lead individual in field chemistry efforts; J. Ameel analyzed much of the water chemistry, and J. Reed helped in the field and laboratory. We thank the EPA–MED laboratory and L. Anderson for analyzing DOC, and for constructive discussions with J. Thompson, J. Morrice, J. Kelly, and R. Kreis. Additional stressor data were developed and summarized by L. Johnson, T. Brown, and G. Host. This research was supported by a grant to G. Niemi from the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) Estuarine and Great Lakes (EaGLe) program through funding to the Great Lakes Environmental Indicators (GLEI) project, U.S. EPA Agreement EPA/R–8286750. This document has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the view of the Agency, and no official endorsement should be inferred. This is contribution number 452 of the Center for Water and the Environment, Natural Resources Research Institute, University of Minnesota Duluth.

Albert, D. A. & Minc, L. D. 2004. Plants as regional indicators of Great Lakes coastal wetland health. *Aquat. Ecosyst. Health Manag.* 7:233–47.

Anderson, N. J. 1990. The biostratigraphy and taxonomy of small *Stephanodiscus* and *Cyclostephanos* species (Bacillariophyceae) in a eutrophic lake, and their ecological implications. *Br. Phycol. J.* 25:217–35.

Anderson, P. & Davic, R. D. 2004. Use of transparency tubes for rapid assessment of TSS and turbidity in streams. *Lake Reservoir. Manag.* 20:110–20.

Bahls, L. L. 1993. *Periphyton Bioassessment Methods for Montana Streams.* Water Quality Bureau, Department of Health and Environmental Sciences, Helena, Montana, 18 pp.

Brazner, J. C., Danz, N. P., Niemi, G. J., Regal, R. R., Trebitz, A. S., Howe, R. W., Hanowski, J. M., et al. 2007. Evaluation of geographic, geomorphic and human influences on Great Lakes wetland indicators: a multi-assemblage approach. *Ecol. Indicators* 7:610–35.

Carlson, R. E. 1977. A trophic state index for lakes. *Limnol. Oceanogr.* 22:361–9.

Charles, D. F. 1996. Use of algae for monitoring rivers in the United States: some examples. *In* Whitton, B. A. & Rott, E. [Eds.] *Use of Algae for Monitoring Rivers II.* Institut für Botanik, AG Hydrobotanik, Universität Innsbruck, Innsbruck, Austria, pp. 109–18.

Chessman, B., Growns, I., Currey, J. & Plunkett-Cole, N. 1999. Predicting diatom communities at the genus level for the rapid biological assessment of rivers. *Freshw. Biol.* 41:317–32.

Danz, N. P., Niemi, G. J., Regal, R. R., Hollenhorst, T., Johnson, L. B., Hanowski, J. M., Axler, R. P., et al. 2007. Integrated gradients of anthropogenic stress in the U.S. Great Lakes Basin. *Environ. Manag.* 39:631–47.

Danz, N. P., Regal, R. R., Niemi, G. J., Brady, V. J., Hollenhorst, T. P., Johnson, L. B., Host, G. E., et al. 2005. Environmentally stratified sampling design for the development of Great Lakes environmental indicators. *Environ. Monit. Assess.* 102:41–65.

Dixit, S. S. & Smol, J. P. 1994. Diatoms as indicators in the Environmental Monitoring and Assessment Program-Surface Waters (EMAP SW). *Environ. Monit. Assess.* 31:275–306.

Environment Canada & U.S. EPA. 2003. *State of the Great Lakes 2003.* EPA 905-R-03-004, U.S. Environmental Protection Agency, Washington, DC, 114 pp.

ESRI. 2000. *ArcGIS Spatial Analyst.* Environmental Systems Research Institute, ESRI Press, Redlands, California.

Fore, L. S. & Grafe, C. 2002. Using diatoms to assess the biological condition of large rivers in Idaho (U.S.A.). *Freshw. Biol.* 47:2015–37.

Fore, L. S., Karr, J. R. & Wisseman, R. W. 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. *J. North Am. Benthol. Soc.* 15:212–31.

Godwin, K. S., Hafner, S. D. & Buff, M. F. 2003. Long-term trends in sodium and chloride in the Mohawk River, New York: the effect of fifty years of road-salt application. *Environ. Pollut.* 124:273–81.

Gray, J. S. 1989. Effects of environmental stress on species rich assemblages. *Biol. J. Linn. Soc.* 37:19–32.

Hall, R. I. & Smol, J. P. 1999. Diatoms as indicators of lake eutrophication. *In* Stoermer, E. F. & Smol, J. P. [Eds.] *The Diatoms: Applications for the Environmental and Earth Sciences.* Cambridge University Press, Cambridge, UK, pp. 128–68.

Hellawell, J. M.1986. *Biological Indicators of Freshwater Pollution and Environmental Management.* Elsevier Applied Science Publishers, New York, 559 pp.

Hill, M. O. & Gauch, H. G. 1980. Detrended correspondence analysis, an improved ordination technique. *Vegetatio* 42:47–58.

Hill, B. H., Herlihy, A. T., Kaufmann, P. R., Stevenson, R. J., McCormick, F. H. & Johnson, C. B. 2000. Use of periphyton assemblage data as an index of biotic integrity. *J. North Am. Benthol. Soc.* 19:50–67.

Host, G. E., Schuldt, J., Ciborowski, J. J. H., Johnson, L. B., Hollenhorst, T. & Richards, C. 2005. Use of GIS and remotely sensed data for a priori identification of reference areas for Great Lakes coastal ecosystems. *Int. J. Remote Sens.* 26:5325–42.

Huston, M. 1979. A general hypothesis of species diversity. *Am. Nat.* 113:81–101.

Juggins, S. 2003. *C2 User Guide. Software for Ecological and Paleoecological Data Analysis and Visualisation.* University of Newcastle, Newcastle upon Tyne, UK, 69 pp.

Karr, J. R. & Chu, E. W. 1999. *Restoring Life in Running Waters: Better Biological Monitoring.* Island Press, Washington, DC, 220 pp.

Karr, J. R., Fausch, K. D., Angermeier, P. L., Yant, P. R. & Schlosser, I. J. 1986. *Assessment of Biological Integrity in Running Waters: A Method and Its Rationale.* Illinois Natural History Survey Special Publication No. 5, Champaign, Illinois, 288 pp.

Kelly, M. G. & Whitton, B. A. 1995. The trophic diatom index: a new index for monitoring eutrophication in rivers. *J. Appl. Phycol.* 7:433–44.

Keough, J. R. & Griffin, J.1994. *Technical Workshop on EMAP Indicators for Great Lakes Coastal Wetlands.* Summary Report. U.S. Environmental Protection Agency and National Biological Service, Duluth, Minnesota, 58 pp.

Kerans, B. L. & Karr, J. R. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. *Ecol. Appl.* 4:768–85.

Kireta, A. R., Reavie, E. D., Axler, R. P., Sgro, G. V., Kingston, J. C., Brown, T. N., Danz, N. P. & Hollenhorst, T. 2007. Coastal geomorphic variability in the Laurentian Great Lakes: implications for a diatom-based monitoring tool. *J. Gt. Lakes Res.* 33:136–53.

Krammer, K. & Lange-Bertalot, H. 1991. Bacillariophyceae. *In* Ettl, H., Gartner, G., Gerloff, J., Hyenig, H. & Mollenhauer, D. [Eds.] *Süsswasserflora von Mitteleuropa 2/4.* Fisher, Stuttgart, Germany, pp. 437.

Lange-Bertalot, H. 1979. Pollution tolerance of diatoms as a criterion for water quality estimation. *Nova Hedwigia* 64:285–304.

Lawson, R. 2004. Coordinating coastal wetlands monitoring in the North American Great Lakes. *Aquat. Ecosyst. Health Manag.* 7:215–21.

Leland, H. V. & Porter, S. D. 2000. Distribution of benthic algae in the upper Illinois River basin in relation to geology and land use. *Freshw. Biol.* 44:279–301.

Lotter, A. F. 2001. The effect of eutrophication on diatom diversity: examples from six Swiss lakes. *In* Jahn, R., Kociolek, J. P., Witkowski, A. & Compère, P. [Eds.] *Lange-Bertalot Festschrift, Studies on Diatoms.* A.R.G. Gantner Verlag K.G., Ruggell, Liechtenstein, pp. 417–32.

Maynard, L. & Wilcox, D. A. 1997. *Coastal Wetlands of the Great Lakes.* State of the Lakes Ecosystem Conference 1996, EPA 905-R-97-015b. Environment Canada and U.S. Environmental Protection Agency, Washington, DC, 70 pp.

MDEQ. 2001. *The Michigan Department of Environmental Quality's Lake Water Quality Assessment Monitoring Program for Michigan's Inland Lakes.* Michigan Department of Environmental Quality, Land and Water Management Division and U.S. Geological Survey, WRD, Michigan District, 4 pp.

Mills, M. R., Beck, G., Brumley, J., Call, S. M., Grubbs, J., Houp, R., Metzmeier, L. & Smathers, K.1993. *Methods for Assessing Biological Integrity of Surface Waters.* Decision of Water, Kentucky Department for Environmental Protection, Frankfort, Kentucky, 189 pp.

MPCA. 2004a. *Citizen Stream Monitoring Program Annual Report for 2003.* Minnesota Pollution Control Agency, Environmental Outcomes Division, St. Paul, Minnesota, 87 pp.

MPCA. 2004b. *Minnesota's Water Quality Monitoring Strategy 2004–2014.* Minnesota Pollution Control Agency, St. Paul, Minnesota, 90 pp.

Mulholland, P. J. 1996. Role in nutrient cycling in streams. *In* Stevenson, R. J., Bothwell, M. L. & Lowe, R. L. [Eds.] *Algal Ecology: Freshwater Benthic Ecosystems.* Academic Press, San Diego, California, pp. 605–39.

Niemi, G. J. & McDonald, M. 2004. Application of ecological indicators. *Annu. Rev. Ecol. Syst.* 35:89–111.

Niemi, G. J., Wardrop, D., Brooks, R., Anderson, S., Brady, V., Paerl, H., Rakocinski, C., Brouwer, M., Levinson, B. & McDonald, M. 2004. Rationale for a new generation of ecological indicators for coastal waters. *Environ. Health Perspect.* 112:979–86.

Nygaard, G. 1949. Hydrobiological studies on some Danish ponds and lakes. II: The quotient hypothesis and some new or little known phytoplankton organisms. *Biol. Skr.* 7:1–293.

Patrick, R. 1973. Use of algae, especially diatoms, in the assessment of water quality. *In* American Society for Testing and Materials [Ed.] *Biological Methods for Assessment of Water Quality.* American Society for Testing Materials, Philadelphia, Pennsylvania, pp. 76–95.

Patrick, R., Hohn, M. H. & Wallace, J. H. 1954. *A New Method for Determining the Pattern of the Diatom Flora.* Academy of Natural Sciences of Philadelphia (No. 259), Philadelphia, Pennsylvania, 12 pp.

Reavie, E. D., Axler, R. P., Sgro, G. V., Danz, N. P., Kingston, J. C., Kireta, A. R., Brown, T. N., Hollenhorst, T. P. & Ferguson, M. J. 2006. Diatom-based weighted-averaging transfer functions for Great Lakes coastal water quality: relationships to watershed characteristics. *J. Gt. Lakes Res.* 32:321–47.

Rosen, B. H. 1995. Use of periphyton in the development of biocriteria. *In* Davis, W. S. & Simon, T. P. [Eds.] *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* Lewis Publishers, Boca Raton, Florida, pp. 209–15.

Round, F. E. & Crawford, R. M. 1990. *The Diatoms. Biology and Morphology of the Genera.* Cambridge University Press, UK, 747 pp.

Schindler, D. W. 1977. Evolution of phosphorus limitation in lakes. *Science* 195:4275.

Seegert, G. 2001. The development, use, and misuse of biocriteria with and emphasis on index of biotic integrity. *Environ. Sci. Pollut.* 3:51–8.

Smol, J. P.2002. *Pollution of Lakes and Rivers: A Paleoenvironmental Perspective.* Arnold Publishers, London, UK, 280 pp.

Stevenson, R. J. 1984. Epilithic and epipelic diatoms in the Sandusky River, with emphasis on species diversity and water pollution. *Hydrobiologia* 114:161–75.

Stevenson, R. J. & Pan, Y. 1999. Assessing ecological conditions in rivers and streams with diatoms. *In* Stoermer, E. F. & Smol, J. P. [Eds.] *The Diatoms: Applications for the Environmental and Earth Sciences.* Cambridge University Press, Cambridge, UK, pp. 11–40.

Stockner, J. C. & Armstrong, F. A. J. 1971. Periphyton of the experimental lakes area, northwestern Ontario. *J. Fish. Res. Board Can.* 28:215–29.

Ter Braak, C. J. F. & Šmilauer, P. 2002. *CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (Version 4.5).* Microcomputer Power, Ithaca, New York, 500 pp.

Tuchman, N. C. 1996. The role of heterotrophy in algae. *In* Stevenson, R. J., Bothwell, M. L. & Lowe, R. L. [Eds.] *Algal Ecology: Freshwater Benthic Ecosystems.* Academic Press, San Diego, California, pp. 299–320.

U.S. EPA. 1999. *EPA Requirements for Quality Assurance Project Plans.* U.S. Environmental Protection Agency, Washington, DC, 40 pp.

Uzarski, D. G., Burton, T. M. & Genet, J. A. 2004. Validation and performance of an invertebrate index of biotic integrity for Lakes Huron and Michigan fringing wetlands during a period of lake level decline. *Aquat. Ecosyst. Health Manag.* 7:269–88.

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.* 28:117–33.

WDNR. 2005. *Water Resources Monitoring Strategy For Wisconsin, Version 1:2-4-2005.* Wisconsin Department of Natural Resources, Madison, Wisconsin, 143 pp.

Whitton, B. A. & Kelley, M. G. 1995. Use of algae and other plants for monitoring rivers. *Aust. J. Ecol.* 20:45–56.

Wilcox, D. A., Meeker, J. E., Hudson, P. L., Armitage, B. J., Black, M. G. & Uzarski, D. G. 2002. Hydrologic variability and the application of Index of Biotic Integrity metrics to wetlands: a Great Lakes evaluation. *Wetlands* 22:588–615.

Williams, J. G. 1964. Possible relationships between plankton-diatom species numbers and water-quality estimates. *Ecology* 45:809–23.

Supplementary Material

The following supplementary material is available for this article:

Appendix S1. Common diatom species and their occurrence statistics and tolerances to various environmental parameters.

Appendix S2. Correlation matrix of candidate metrics.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/ 10.1111/j.1529-8817.2008.00523.x.

(This link will take you to the article abstract.)

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.