J-NABS

Journal of the
North American
Benthological
Society

A weight-of-evidence approach to define nutrient criteria protective of aquatic life in large rivers

Alexander J. Smith[1]

Stream Biomonitoring Unit, New York State Department of Environmental Conservation, 425 Jordan Road, Troy, New York 12180 USA

Christopher P. Tran[2]

New England Interstate Water Pollution Control Commission, Boott Mills South, 116 John Street, Lowell, Massachusetts 01852 USA

J. N. Am. Benthol. Soc., 2010, 29(3):875–891
© 2010 by The North American Benthological Society
DOI: 10.1899/09-076.1
Published online: 8 June 2010

A weight-of-evidence approach to define nutrient criteria protective of aquatic life in large rivers

Alexander J. Smith[1]
Stream Biomonitoring Unit, New York State Department of Environmental Conservation, 425 Jordan Road, Troy, New York 12180 USA

Christopher P. Tran[2]
New England Interstate Water Pollution Control Commission, Boott Mills South, 116 John Street, Lowell, Massachusetts 01852 USA

Abstract. Cultural eutrophication of surface waters has become a major source of water-quality impairment throughout the US. In response, the US Environmental Protection Agency (USEPA) has devised a national strategy for the development of regional nutrient criteria. Our study is part of New York State's effort to revise its narrative nutrient standard for N and P and is based on the USEPA's recommended weight-of-evidence approach. The objective of our investigation was to identify nutrient thresholds based on a final weighted average of results from percentile analysis, nonparametric deviance reduction (changepoint), and cluster analysis. The thresholds were determined from shifts in biological community structure (benthic macroinvertebrate and diatom) related to water-column nutrient data from 40 large river sites throughout New York State. USEPA's percentile analysis yielded possible criteria of 0.023 mg total P (TP)/L, 0.51 mg total N (TN)/L, 0.16 mg NO_3-N /L, and 2.4 mg chlorophyll *a* (chl *a*)/m^3. Threshold responses in benthic macroinvertebrate metrics at the 50th percentile occurred at concentrations between 0.009 and 0.07 mg TP/L, 0.41 and 1.2 mg TN/L, 0.18 and 0.55 mg NO_3-N/L, and 2.1 mg chl *a*/m^3. Cluster analysis yielded 3 groups of sites based on macroinvertebrate and diatom taxa. The median nutrient values of the medium-nutrient-condition site clusters were used to set criteria for TP and TN. For site clusters based on macroinvertebrate data these values were 0.037 mg TP/L and 0.68 mg TN/L. For clusters based on diatom data these were 0.037 mg TP/L and 0.78 mg TN/L. Based on the weight-of-evidence approach and results from all 3 methods, the proposed guidance values for nutrients in large rivers are 0.03 mg TP/L, 0.7 mg TN/L, 0.3 mg NO_3-N/L, and 2.2 mg chl *a*/m^3. These values are similar to those derived by others and provide meaningful nutrient endpoints that would be protective of aquatic life in large rivers.

Key words: macroinvertebrates, diatoms, nutrient criteria, nutrients, nonparametric deviance reduction, cluster analysis, percentile analysis.

Anthropogenic inputs of N and P have become dominant causes of water-quality impairment in the US (USEPA 2000d, Davis 2002, Reckhow et al. 2005, Smith et al. 2007). Thus, eutrophication of surface waters has become an important reason to develop effects-based water-quality standards for aquatic life use. In accordance with the US Environmental Protection Agency's (USEPA) national strategy for the development of nutrient criteria (USEPA 1998), states like New York have begun to establish their own nutrient criteria goals rather than adopt USEPA guidance values. The economic consequences of controlling nutrients in surface water and the extent to which nutrients have affected them mean that such goals must be based on the most credible and defensible scientific data (Dodds and Welch 2000, Paul and McDonald 2005).

The reasons for needing nutrient criteria and the methods for deriving them are diverse. Eutrophication of downstream systems from upstream impairments, human-health effects of excessive nutrient loads in water supplies, effects on recreational use and aesthetics, and impact on aquatic biota are

[1] E-mail address: ajsmith@gw.dec.state.ny.us

[2] Present address: State University of New York at Ulster, 71 Smith Avenue, Kingston, New York 12401 USA. E-mail: tranc@dep.nyc.gov

driving forces behind the national call for nutrient criteria (Dodds and Welch 2000). Setting criteria to protect against these conditions is challenging. Some researchers have determined regional background nutrient conditions and applied them as guidance (USEPA 2000a, b, c, 2001, Rohm et al. 2002), but these criteria lack direct association with many of the biological communities they are meant to protect. Some researchers have developed databases of nutrient information (Palmstrom 2002) and have used the USEPA's method of reference and test populations to define criteria (Martinez 2002). Some have devised classification and modeling approaches (Snelder et al. 2004) or developed rating curves and load estimates using water-chemistry data (Sheeder and Evans 2004). Others have taken more predictive approaches (Reckhow et al. 2005) and have tried to relate biotic assemblages to threshold levels of nutrients (Havens 2003, Smith et al. 2007, Wang et al. 2007).

Recent efforts have been directed toward identifying nutrient thresholds that represent significant shifts in aquatic ecosystem structure and function (Dodds et al. 1997, 1998, Dodds and Welch 2000, Havens 2003, Sheeder and Evans 2004, Stevenson et al. 2006, Smith et al. 2007, Wang et al. 2007). The *weight-of-evidence* approach has been recommended by the USEPA to address the complex nature of deriving nutrient criteria. This approach combines the results from several methods to describe the nutrient–biota relationship in the aquatic environment. After determining the level of confidence placed on results from different analyses, professional judgment is used to weight the endpoints. The resulting criteria tend to balance the significance of the results with the appropriate levels of protection for aquatic life. However, use of professional judgment mandates that the procedure of applying weights to individual results be transparent when defining the final criteria (USEPA 2000e, Tetra Tech 2008).

Our study is part of a larger effort by New York State (NYS) to revise its narrative nutrient standard for N and P that reads "none in amounts that will result in growths of algae, weeds and slimes that will impair the waters for their best usages" (NYS Environmental Conservation Law Chapter X, Parts 700–706). Similar work already has been completed for wadeable streams in NYS (Smith et al. 2007). The objective of our investigation was to use the USEPA's recommended weight-of-evidence approach to identify nutrient thresholds associated with shifts in benthic macroinvertebrate and periphyton community structure in large, nonwadeable river systems in NYS and to provide an example for other states to follow. Our results could then be used in replacing NYS's narrative standard with scientifically supported numeric values. The investigation consisted of 2 major parts: 1) to determine effects of varying nutrient concentrations on algal and invertebrate community structure in large rivers and 2) to propose protective numeric nutrient thresholds based on the information gathered.

Methods

Study design and sample collection

We needed to establish a set of physical criteria that would define large rivers as nonwadeable flowing waters with greater influence from autotrophic and planktonic production and longer retention times than might be encountered in wadeable streams. Historical NYS Department of Environmental Conservation (DEC) water-quality sampling locations and variables used by other states were reviewed. In Idaho, stream width (>30 m), depth (>0.4 m), and order ($\geq$4) are used to differentiate between wadeable streams and rivers (Fore and Grafe 2002). In Ohio, the boundary between streams and rivers is defined solely on drainage area (>518 km^2) (Miltner and Rankin 1998). The USEPA includes Strahler stream order, whether the system is boatable, and dominance of riverine species as other attributes for classifying large rivers (Flotemersch et al. 2006). The USEPA also uses a comprehensive definition of large rivers based on drainage area ($\geq$1600 km^2) and depth ($\geq$1 m) (Wilhelm et al. 2005). In NYS, DEC sampling locations >1 m deep typically occur on streams with drainage area $\geq\sim$1295 km^2. Therefore, we defined large river sites as those with a drainage area of $\geq$1295 km^2 and depth $\geq$1 m. During field reconnaissance, we found that some sites that met these criteria were very similar to wadeable streams because of the presence of high-gradient riffle areas. Therefore, we revised the definition to include a requirement that the region of the river upstream of the sampling site must be free of riffles for a distance of 20$\times$ the wetted width of the river.

We selected 40 sites from a pool of ~11,000 possible sampling sites in NYS in the US Geological Survey's (USGS) National Water Inventory System (NWIS). We viewed NWIS sites as a suitable starting pool of possible sampling sites because: 1) each station has a delineated watershed with known drainage area, 2) background information for some type of environmental variable (e.g., discharge, water chemistry, biology) existed for each site; and 3) the number of sites and their distribution provided a good representation of large rivers across NYS. Of the 11,000 NWIS sites, ~7790 had a drainage area $\geq$1295 km^2.

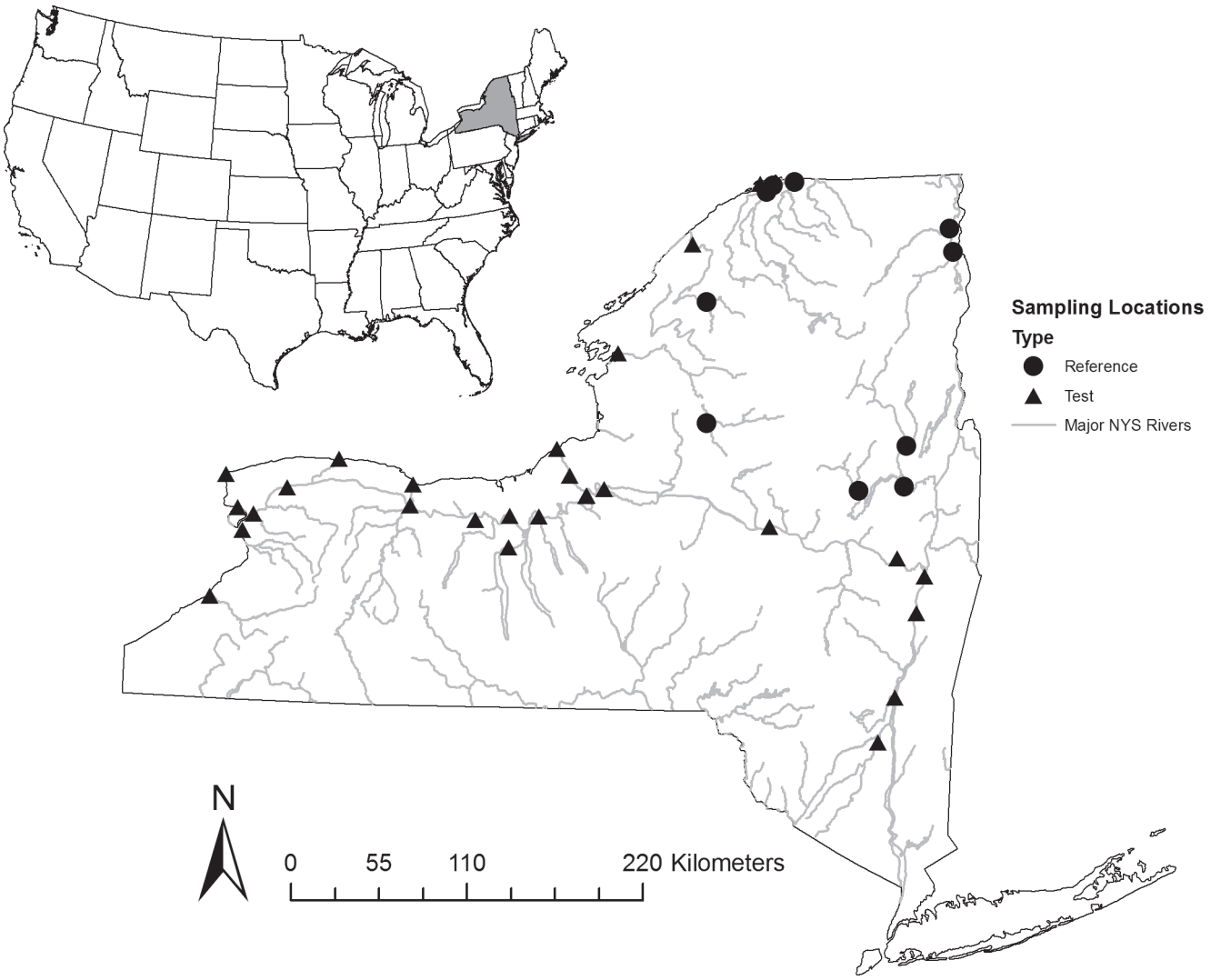

FIG. 1. Reference and test site sampling stations on large rivers in New York. Most reference sites were in aggregate nutrient ecoregion VII (mostly glaciated dairy) but drained the minimally disturbed ecoregion VIII (nutrient-poor, largely glaciated upper midwest and northeast). NYS = New York State.

We selected a set of 10 reference and 30 test sites from the candidate pool of NWIS sites (Fig. 1). We selected reference sites based on a set of predefined criteria to ensure that they represented least-disturbed, best-attainable conditions from throughout the state (Reynoldson et al. 1997). The selection criteria consisted of biotic and abiotic factors, including watershed-scale land use >75% natural cover, 'unimpacted' historical biological condition assessments, water-chemistry values (when available) within NYS background conditions (Parsons and Norris 1996, Maxted et al. 2000), and best professional judgment (Hughes 1995).

We sampled sites between July and September 2006 and 2007. In 2006, we visited the original 40 sites, but found that 15 of these sites were similar to wadeable stream sites. We removed these sites from the study (Hughes 1995), revised our definition of large rivers, and selected 15 new sites for sampling in 2007.

The final set of sites (Fig. 1) provided a data set covering a large area of NYS and had similar biological communities, water chemistry, and physical attributes. Seven reference sites were in Ecoregion VII (mostly glaciated dairy region), one of 14 nutrient ecoregions developed for establishing nutrient criteria throughout the US. These ecoregions were selected for nutrient criteria based on similarities among many landscape-level features including biogeography (USEPA 2000e). Three reference sites were at the boundary of Ecoregion VII and Ecoregion XI (central

and eastern forested uplands) and represented a different ecoregion. Our characterization of reference conditions was dominated by the 7 sites in Ecoregion VII, but the 3 remaining sites facilitated adequate comparisons in the data set by including natural community variation.

We sampled benthic macroinvertebrates and periphyton with modified Hester–Dendy multiplate artificial substrate samplers, each consisting of 3 tempered hardboard plates (Hester and Dendy 1962, Smith et al. 2009). Multiplates were positioned ~1 m below the water surface and secured by flotation devices and anchor weights to prevent movement. We made an effort to ensure consistency among sites by placing samplers in main channel flows and away from obvious human disturbances, such as boat launches or discharges. Samplers were deployed for a colonization period of 5 wk. We deployed 3 samplers at each station although only 2 were needed for analysis (1 for benthic macroinvertebrate community analysis and 1 for periphyton). The extra sampler increased the probability of retrieving at least 2 multiplates. We scraped all organisms and other material that had colonized a sampler into a bucket (1 for each sampler). We poured the material through US no. 30 standard sieve (600 μm). We rinsed the sample from the sieve with water from the site into a 118-mL glass jar. Macroinvertebrate samples were preserved in 95% ethyl alcohol and periphyton samples in 4% formaldehyde.

Macroinvertebrate samples were processed by the NYS Department of Environmental Conservation (NYSDEC) Stream Biomonitoring Unit (SBU) with its standard method (Smith et al. 2009). Samples were rinsed with tap water in a US no. 40 standard sieve (425 μm) and subsampled by placing the sample in a tray, evenly distributing it over the bottom, and using a divider to split the sample into quarters. All organisms >1.5 mm long were removed from 1 quarter until 250 individuals had been sorted or the quarter was finished. Additional quarters were not sorted. Macroinvertebrates were identified to lowest-possible taxonomic resolution.

We sent periphyton samples to a contract laboratory for processing and identification of 300 cells. Samples were digested with heat and acid washing by the procedures outlined in Smith et al. (2009). Deionized water was added to the digested sample. The sample was allowed to settle and was decanted, and fresh deionized water was added. This process was repeated until a neutral pH was achieved. Slides were made for each sample by arranging cover slips on a slide warmer and adding the clean diatom sample to the cover slip with a pipette. After drying, cover slips were examined at 400× magnification. Cover slips were permanently mounted on labeled slides with a high-resolution mounting medium. Diatoms were counted and identified to the lowest practical taxon until ≥300 cells (600 valves) were encountered. Counting and identification was done at 1000× magnification.

We collected samples for water-chemistry analysis twice at each sampling site (at the time of multiplate deployment and at the time of retrieval). We collected samples with a DH-81 depth-integrating sampler (Rickly Hydrological Company, Columbus, Ohio) attached to a 1-m rod. The samples integrated the top 1 m of surface water. We sent samples to a contract laboratory for analysis of primary nutrient variables recommended for nutrient criteria development by the USEPA (USEPA 2000e). Analyses were performed using standard methods (USEPA 1993): total P (TP; detection limit = 0.003 mg/L, method 365.1), NO_3^--N (detection limit = 0.05 mg/L, method 353.2), total N (NH_4^+ [method 350.1R] + NO_3^-/NO_2^--N + total organic N [method 351.5]), chlorophyll *a* (chl *a*; detection limit = 2.0 mg/m^3, method 1002G), and turbidity (Tb; detection limit = 0.1 NTU, method 180.1). Results of the 2 samples from each station were combined to form station means. Several other water-chemistry variables (NH_4^+, Cl^-, NO_3^- + NO_2^--N, NO_2^--N, PO_4^{3-}, Si, SO_4^{2-}, total alkalinity, total Kjeldahl N, and total organic N) were analyzed to provide supporting information when needed, but they will not be discussed further.

Statistical analyses

Percentile analysis.—We used a percentile analysis of the water-chemistry data to determine possible criteria for the variables TP, TN, NO_3-N, chl *a*, and Tb (USEPA 2000e). In its technical guidance for developing numeric nutrient criteria, the USEPA suggests using nutrient concentrations at the 75th percentile of reference sites as criteria. USEPA also recommends use of the 5th to 25th percentile of the entire population, especially where reference sites are not established (USEPA 2000e). We used a modification of this method that incorporated reference- and test-site populations separately to retain the information contained in both data sets. We determined the 75th percentile of the reference condition and the 25th percentile of the test condition and used the median value of the 2 percentiles for each response variable as the potential nutrient criteria for TP, TN, NO_3-N, chl *a*, and Tb.

Percentile analysis is sufficient according to USEPA guidance, but the criteria derived by this method are not linked to water body designated uses, such as

aquatic life use. Nutrient gradients must be related to biological outcomes to develop meaningful nutrient criteria within a regulatory framework. We used Spearman Rank Order correlation to identify relationships between biological condition (community and water-quality metrics) and primary nutrient criteria variables (USEPA 2000e). We plotted and evaluated each relationship to ensure that weak correlations were not simply the result of nonlinear relationships between independent and response variables. We used Bonferroni corrections to reduce the risk of making Type I errors because of the large number of comparisons. Therefore, we set $\alpha = 0.006$ for correlations between nutrients and macroinvertebrate metrics and $\alpha = 0.007$ for correlations between nutrients and diatom metrics.

We analyzed a standard set of benthic macroinvertebrate community metrics used by NYS to assess water quality and nutrient enrichment in large rivers. The metrics were species richness, Ephemeroptera, Plecoptera, and Trichoptera (EPT) richness (Lenat 1988), Hilsenhoff's biotic index (HBI) score (Hilsenhoff 1987), Shannon–Wiener diversity (DIV), non-Chironomidae and Oligochaeta (NCO) richness (Riva-Murray et al. 2002, Smith et al. 2009), NYS's biological assessment profile (BAP) score, and Nutrient Biotic Indices for TP (NBI-P) and NO_3-N (NBI-N) (Smith et al. 2007).

NYS does not currently have a standard set of periphyton community metrics to use when assessing water quality and nutrient enrichment in large rivers. However, one is under development, so we tested the periphyton community metrics that NYS is currently evaluating. These metrics were species richness; Shannon–Wiener diversity (log[x]-transformed; Magurran 2004); Van Dam's Trophic State Index (TRI) (Van Dam et al. 1994); % oligotrophic, % mesotrophic, and % eutrophic individuals (Van Dam et al. 1994); and a Pollution Tolerance Index (PTI) (Lange-Bertalot 1979).

Changepoint analysis.—To set meaningful criteria, specific values of environmental variables must be identified that cause or reflect a change in biological community structure. We used nonparametric changepoint analysis (nCPA) (King and Richardson 2003, Qian et al. 2003, 2004, King et al. 2005, 2007) to identify thresholds in the biological response to increasing concentrations of nutrients. nCPA selects the point along the independent-variable gradient that produces the greatest reduction in deviance. The changepoint can be any value that creates 2 populations in the data set that are separated by significant differences in their mean or variance (King and Richardson 2003, Qian et al. 2003, King et al. 2005). We ran nCPA with the custom function chngp.nonpar (King and Richardson 2003, Qian et al. 2003) in S-PLUS 6.1 Professional (Insightful Corporation, Seattle, Washington).

We used nCPA to test for threshold responses of community metrics that were significantly correlated with nutrient criteria (Spearman rank order). Changepoint values have associated uncertainty because any one value or multiple values of the predictor variable could potentially be a changepoint. nCPA uses bootstrap resampling with replacement (1000 permutations) to estimate uncertainty in the changepoint values and produces cumulative probability plots for each comparison based on the frequency distribution of changepoints. The method also provides a description of the confidence in each changepoint because of the nature of the resampling (King and Richardson 2003). The ability to create cumulative probability plots from the results of the nCPA analysis affords a degree of risk assessment in the changepoints (King and Richardson 2003) and can be useful when selecting final nutrient criteria.

Cluster analysis.—Neither the USEPA percentile method nor nCPA make full use of information provided by natural community shifts in species composition along nutrient gradients. Cluster analyses provide insight into how community composition shifts as nutrient concentrations change. We used Bray–Curtis similarity analysis (BCA) (Bray and Curtis 1957) to identify clusters of similar sites based on log(x)-transformed macroinvertebrate and diatom species data (Smith et al. 2007). We used dendrograms to identify groups of sites with greater species composition similarity and dissimilarities between groups (Simon and Morris 2009). We used Kruskal–Wallis 1-way analyses to determine if nutrient concentrations differed among the clusters identified by BCA. In these tests, we used nutrient concentrations as response variables and clusters of sites as categorical variables. We used Tukey's multiple comparison tests to determine which clusters differed.

Results

Percentile analysis

Based on percentile analysis (median value of the 75th percentile of the reference sites and the 25th percentile of the test sites), numeric nutrient criteria would be: 0.023 mg TP/L, 0.51 mg TN/L, 0.16 mg NO_3-N/L, 2.4 mg chl a/m^3, and Tb = 3.0 NTU (Table 1). These values are similar to those proposed in other studies (Dodds et al. 1997, 1998), and they are almost identical to USEPA nutrient criteria (USEPA 2000a).

Spearman rank order correlation results suggest significant relationships between primary nutrient

TABLE 1. Results of percentile analysis conducted on primary nutrient criteria variables showing the 75th percentile of the variable at sites in reference condition and 25th percentile of the variable at test sites. Potential nutrient criteria are given as the median value of the 2 percentiles for each variable.

Response variable	Reference site 75th percentile (n = 10)	Test site 25th percentile (n = 30)	Median
Total P (mg/L)	0.018	0.029	0.023
Total N (mg/L)	0.48	0.53	0.51
NO_3-N (mg/L)	0.15	0.17	0.16
Chlorophyll *a* (mg/m^3)	2.4	2.3	2.4
Turbidity (NTU)	3.3	2.7	3.0

criteria variables and 3 macroinvertebrate community metrics (Table 2). The NBI-P, a benthic macroinvertebrate metric developed specifically to address nutrient enrichment from TP, was significantly correlated with TP, TN, and NO_3-N. However, the NBI-N, a similar metric to the NBI-P but for NO_3^-, was not significantly correlated with any of the nutrient variables (Table 2). HBI was the only macroinvertebrate metric that was correlated with chl *a*. The BAP, the NYS multimetric index for assessment of water quality, was correlated only with TP. None of the metrics evaluated was significantly correlated with Tb.

Diatom community metrics were more strongly correlated with nutrients than were macroinvertebrate community metrics (Table 3). Diatom metrics with the strongest correlations to water chemistry were those associated with site trophic status, specifically % eutrophic and % mesotrophic individuals. None of the diatom metrics was significantly correlated with chl *a*, possibly because of the high detection limit associated with the chl *a* analysis used and the relatively low concentration of chl *a* in NYS waters. Percent eutrophic was the only diatom metric correlated with Tb. PTI, but not TRI, was significantly correlated with several nutrients (TN and NO^3-N).

Changepoint analysis

Only metrics that were significantly correlated with primary nutrient criteria variables were used in the nCPA. Threshold responses (changepoints) were found for several biological community metrics at various concentrations of most primary nutrient criteria variables (Table 4). We used only those changepoints with the lowest associated Type I errors in further analyses. NBI-P (macroinvertebrates), % mesotrophic (diatoms), and % eutrophic (diatoms) had changepoints along the TP gradient (Table 4; Fig. 2A–C). These metrics specifically address nutrient enrichment. NBI-P, % mesotrophic, % eutrophic, and PTI had changepoints along the TN gradient (Fig. 3A–C). NBI-P (Fig. 4) and PTI had changepoints along the NO_3-N gradient, and HBI had a changepoint along the chl *a* gradient (Table 4).

Cumulative probability distributions suggested that threshold responses of biological community metrics occurred between 0.009 and 0.07 mg TP/L, 0.41 and 1.22 mg TN/L, 0.18 and 0.55 mg NO_3-N/L, and at 2.1 mg chl *a*/m^3. The 50th percentile threshold for a changepoint in BAP along a TP gradient was 0.07 mg/L. The mean BAP score at sites with TP $<$ 0.07 mg/L was 7.5, whereas the mean BAP score for sites with TP $>$ 0.07 mg/L was 5.0. These scores correspond to 'slightly impacted' and 'moderately impacted', respectively, assessments in the NYS tiered assessment framework (Riva-Murray et al. 2002, Smith et al. 2009) (Table 4). The BAP was not correlated with TN and NO_3-N, so these relationships were not evaluated

TABLE 2. Spearman rank order correlation results for benthic macroinvertebrate community metrics and primary nutrient criteria variables. HBI = Hilsenhoff's biotic index score, EPT = Ephemeroptera, Plecoptera, and Trichoptera richness, DIV = Shannon–Wiener diversity, NCO = non-Chironomidae and Oligochaeta richness, BAP = Biological Assessment Profile score, NBI-P = Nutrient Biotic Index-P, NBI-N = Nutrient Biotic Index-NO_3, TP = total P, TN = total N, chl *a* = chlorophyll *a*, Tb = turbidity. * indicates significant relationships at $p \leq 0.006$ (Bonferroni correction for multiple statistical tests).

Nutrient	Statistic	Richness	HBI	EPT	DIV	NCO	BAP	NBI-P	NBI-N
TP	r	−0.279	0.510	−0.375	−0.364	−0.355	−0.445	0.504	0.424
	p	0.094	0.001*	0.022	0.027	0.031	0.006*	0.002*	0.009
TN	r	−0.190	0.411	−0.233	−0.312	−0.268	−0.320	0.504	0.424
	p	0.260	0.012	0.163	0.060	0.108	0.054	0.002*	0.009
NO_3	r	−0.105	0.086	−0.057	−0.188	−0.176	−0.116	0.566	0.344
	p	0.535	0.610	0.735	0.263	0.296	0.490	0.000*	0.038
Chl *a*	r	−0.153	0.468	−0.294	−0.117	−0.242	−0.251	0.262	0.194
	p	0.362	0.004*	0.077	0.488	0.147	0.134	0.116	0.248
Tb	r	−0.005	0.246	−0.120	−0.019	−0.126	−0.083	0.299	0.304
	p	0.977	0.140	0.477	0.908	0.455	0.625	0.072	0.067

TABLE 3. Spearman rank order correlation results for diatom community metrics and primary nutrient criteria variables. DIV_D = log(Shannon–Wiener diversity), TRI = Van Dam's Trophic State Index, % oligotrophic = % oligotrophic individuals, % mesotrophic = % mesotrophic individuals, % eutrophic = % eutrophic individuals, PTI = Pollution Tolerance Index, TP = total P, TN = total N, chl *a* = chlorophyll *a*, Tb = turbidity. * indicates significant relationships at $p \leq 0.007$ (Bonferroni correction for multiple statistical tests).

Variable	Statistic	Richness	DIV_D	TRI	% oligotrophic	% mesotrophic	% eutrophic	PTI
TP	*r*	−0.194	−0.208	0.183	−0.373	−0.643	0.714	−0.388
	p	0.248	0.216	0.277	0.023	0.000*	0.000*	0.018
TN	*r*	−0.051	−0.079	0.134	−0.306	−0.556	0.655	−0.489
	p	0.762	0.640	0.429	0.065	0.000*	0.000*	0.002*
NO_3	*r*	0.044	0.084	0.105	−0.176	−0.320	0.329	−0.500
	p	0.793	0.618	0.536	0.294	0.054	0.047	0.002*
Chl *a*	*r*	−0.033	−0.002	−0.073	−0.220	−0.294	0.357	−0.249
	p	0.847	0.992	0.664	0.189	0.077	0.030	0.136
Tb	*r*	0.189	0.181	−0.114	−0.388	−0.404	0.604	−0.407
	p	0.260	0.283	0.499	0.018	0.013	0.000*	0.013

with nCPA. The 50th percentile threshold for a changepoint in NBI-P was 0.011 mg TP/L. The mean NBI-P score at sites with TP < 0.011 was 5.38 (mesotrophic, borderline oligotrophic) (Smith et al. 2007), whereas the mean NBI-P score at sites with TP > 0.011 mg TP/L was 7.46 (eutrophic) (Table 4, Fig. 2). NBI-P scores did not respond as strongly to gradients of TN and NO_3-N as to TP (Table 4, Figs 3A, 4A).

Cluster analysis

BCA identified differences in macroinvertebrate and diatom species composition of sites along nutrient gradients (Fig. 5A, B). The dendrograms show 3 major clusters for both macroinvertebrates and periphyton at ~30 to 40% similarity. At higher levels of similarity (>40%) additional clusters are apparent. We used the clusters identified at 30 to 40% similarity for the purpose of establishing nutrient criteria. Clusters corresponded to a group of reference sites (reference-site cluster) and 2 groups of test sites (test-site clusters 1 and 2) (Fig. 5A, B).

TP differed significantly between the macroinvertebrate reference-site cluster and test-site cluster 2 and between the test-site clusters 1 and 2 (Fig. 6A). Median values of TP associated with the macroinvertebrate reference-site cluster and test-site clusters 1 and 2 were 0.011 mg/L, 0.037 mg/L, and 0.07 mg/L, respectively. Similar results were obtained with diatom-based clusters, but TP values did not differ significantly between the 2 test-site clusters (Fig. 6B). Median values of TP associated with the diatom reference-site cluster and test-site clusters 1 and 2 were 0.011 mg/L, 0.037 mg/L, and 0.06 mg/L respectively.

TN differed between the macroinvertebrate reference-site cluster and the 2 test-site clusters but not between test-site clusters 1 and 2 (Fig. 7A). Median values of TN for macroinvertebrate clusters differed from median values for corresponding diatom clusters (macroinvertebrates: reference-site = 0.45mg/L, test-site 1 = 0.68mg/L, test-site 2 = 1.06 mg/L; diatoms: reference-site = 0.44mg/L, test-site 1 = 0.77mg/L, test-site 2 = 1.19mg/L). NO_3-N and chl *a* did not differ significantly among macroinvertebrate or diatom clusters.

Weight-of-evidence nutrient criteria

We combined the results of each line of evidence in a weighted mean to define final nutrient criteria. Weights were assigned on the basis of strength and significance of the analysis, confidence in the data, and best professional judgment (BPJ). The multiple lines of evidence used were percentiles, metrics that yielded significant changepoints, and cluster analyses. Results from metrics established specifically for or directly related to nutrients in the water were weighted more heavily than those associated with general pollution or, in the case of the percentile analysis, had no connection with biological responses. To reduce the subjectivity of applying weights to individual results and to provide a more reproducible method, analyses were placed on a scale of increasing connection with biological response to nutrients. The scale ranged from 1 to 2. Analyses with no connection to biological assemblages (percentiles) received a weight of 1, analyses that provided an indirect evaluation of response to nutrients (BCA) received a weight of 1.5, and analyses of direct or threshold responses to gradients (nCPA) received a weight of 2.

For TP, we used (weights applied are given in parentheses) the changepoints associated with the

TABLE 4. Results from nonparametric changepoint analysis using only biological water-quality metrics that were significantly correlated with primary nutrient criteria variables. Threshold responses for the 5th, 50th, and 95th percentile of cumulative probabilities are given. Probability of Type I error (p) is for 50th percentiles only. Mean scores at sites with nutrient concentrations below (left) or above (right) the 50th percentile are provided for each water-quality metric. BAP = Biological Assessment Profile Score, NBI-P = Nutrient Biotic Index-P, % mesotrophic = % mesotrophic individuals, % eutrophic = % eutrophic individuals, PTI = Pollution Tolerance Index, TP = total P, TN = total N, chl *a* = chlorophyll *a*, Tb = turbidity. The α value for p was established using Bonferroni correction for multiple statistical tests and varied depending on nutrient variable and number of tests performed. * indicates the test was significant, these values were used in defining the final proposed criteria.

Primary nutrient criterion variable	Water-quality metric	Changepoint (percentiles of cumulative probability)			p	Mean metric score	
		5th	50th	95th		Left	Right
TP (mg/L) $\alpha = 0.01$	BAP	0.028	0.070	0.143	0.0284	7.5	5
	NBI-P	0.011	0.011	0.036	0.0026*	5.38	7.46
	HBI	0.009	0.030	0.143	0.0220	5.97	7.1
	% mesotrophic	0.009	0.009	0.013	0.0006*	30	4
	% eutrophic	0.011	0.020	0.077	0.0010*	27	65
TN (mg/L) $\alpha = 0.01$	NBI-P	0.44	0.51	0.76	0.0061*	5.91	7.59
	% mesotrophic	0.39	0.41	0.48	0.0072*	25	5
	% eutrophic	0.45	0.50	1.14	0.0019*	30	65
	PTI	1.03	1.22	1.28	0.0004*	2.67	2.25
NO_3-N (mg/L) $\alpha = 0.03$	NBI-P	0.15	0.18	0.27	0.0021*	5.97	7.79
	PTI	0.15	0.55	0.65	0.0024*	2.66	2.30
Chl *a* (mg/m^3) $\alpha = 0.05$	HBI	2.1	2.1	21.2	0.0259*	5.32	6.93

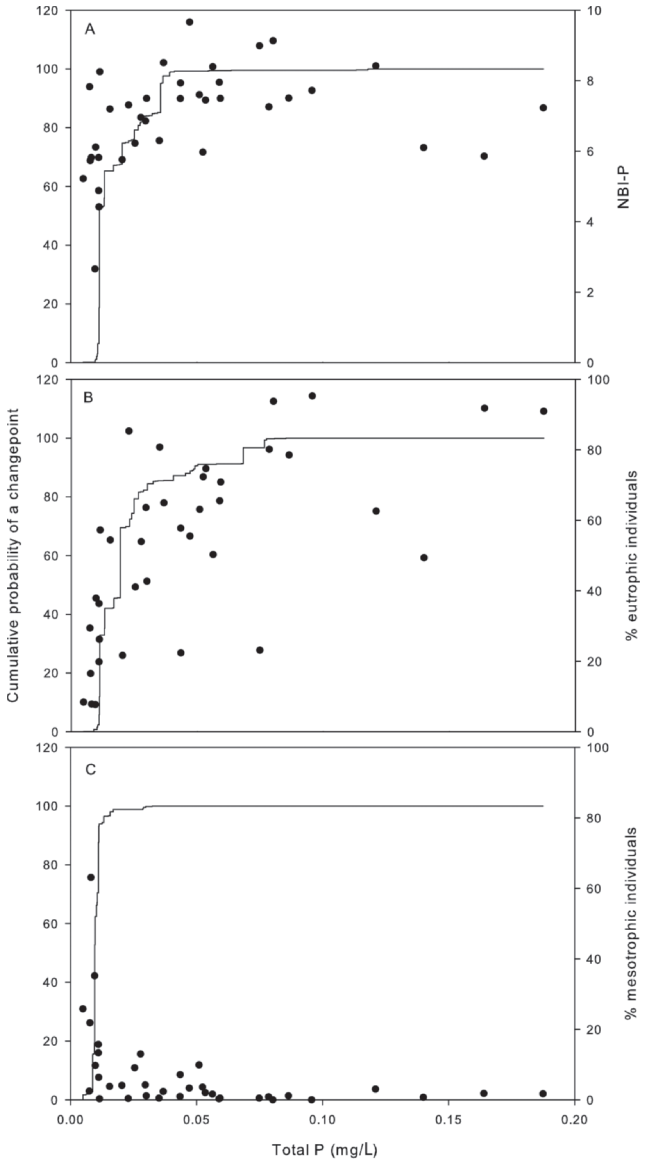

FIG. 2. Cumulative probability plots for significant changepoints of the Nutrient Biotic Index-P (NBI-P; based on benthic macroinvertebrates) (A), % eutrophic diatom individuals (B), and % mesotrophic diatom individuals (C) to increasing total P (TP). Lines represent the cumulative probability that a changepoint occurs at a given TP concentration. Points represent raw metric scores at sites along the TP gradient.

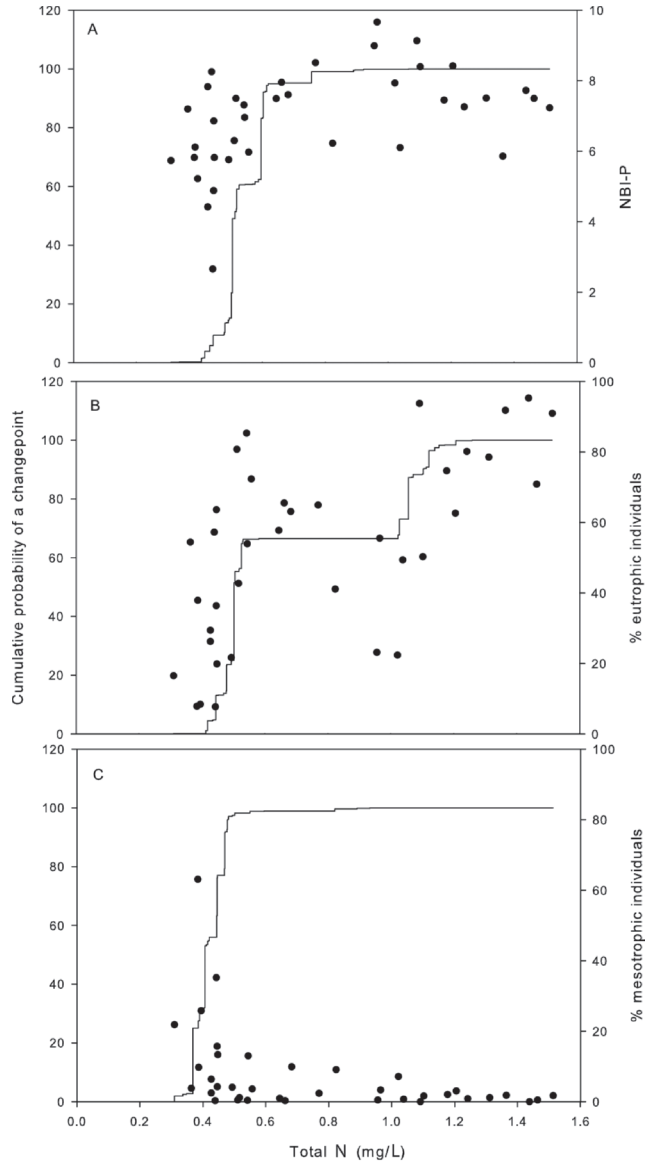

FIG. 3. Cumulative probability plots for significant changepoints of the Nutrient Biotic Index-P (NBI-P; based on benthic macroinvertebrates) (A), % eutrophic diatom individuals (B), and % mesotrophic diatom individuals (C) to increasing total N (TN). Lines represent the cumulative probability that a changepoint occurs at a given TN concentration. Points represent raw metric scores at sites along the TN gradient.

50th percentile for NBI-P (2), % mesotrophic (2), % eutrophic (2), and the BAP (2). The BAP was retained in calculating the final value for TP because it was significantly correlated with TP and is important for linking nutrient criteria to NYS water-quality assessment methods. We also used the median value associated with the percentile analysis (1) and the median values associated with the macroinvertebrate (1.5) and diatom (1.5) test-site 1 (medium nutrient condition) clusters in the BCA analysis.

TN, NO_3-N, and chl *a* values were defined using the same method. However, the metrics used changed according to the relationship with the nutrient variable. For TN, we used NBI-P (2), % mesotrophic (2), % eutrophic (2), PTI (2), percentile analysis (1), and cluster analyses (1.5). The criterion for NO_3-N

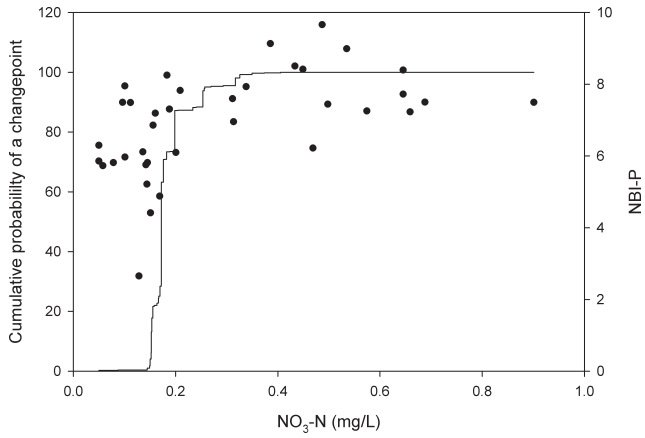

FIG. 4. Cumulative probability plot for a significant changepoint of the Nutrient Biotic Index-P (NBI-P) to increasing NO_3-N. The line represents the cumulative probability that a changepoint occurred at a given NO_3-N concentration. Points represent the raw metric scores at sites along the NO_3-N gradient.

was based on fewer metrics because fewer metrics were significantly correlated with NO_3-N. These metrics were the NBI-P (2), PTI (2), and percentile analysis (1). Even fewer metrics were significantly correlated with chl *a*, and only the HBI (2) and percentile analysis (1) were used to define the proposed criterion. More confidence should be placed in the TP and TN criteria because these values were supported by the most lines of evidence. Based on the weight-of-evidence approach, the proposed large river nutrient criteria are: 0.03 mg TP/L, 0.7 mg TN/L, 0.3 mg NO_3-N/L, and 2.2 mg chl *a* /m^3. Results for Tb were inconclusive, and no values are proposed.

Discussion

Percentile analysis

We used multiple lines of evidence to identify nutrient concentration thresholds related to shifts in biological community structure. Evaluation of percentile, changepoint, and cluster analyses resulted in nutrient criteria directly related to aquatic life uses. Each analysis provided independent values for TP, TN, NO_3-N, and chl *a*. All of these values are arguably reasonable regulatory goals for nutrient conditions in large rivers. However, combining these values would generate nutrient criteria that are more closely related to the effects of nutrients on aquatic life.

The USEPA method of using percentiles to define nutrient criteria provides only a means of examining nutrient concentration frequency in a given population of sites and ignores the responses of biological communities to increases in nutrients. The diversity of possible population classes and the natural variability of nutrients throughout the nation justified USEPA's ecoregional approach to using percentiles (USEPA 2000e, Martinez 2002, Rohm et al. 2002). However, the percentile-based guidance values set forth by USEPA tend to be significantly lower than independently derived criteria (Ice and Binkley 2003, Smith et al. 2007) or criteria already in use by regulatory agencies (Martinez 2002). For example, Ice and Binkley (2003) reviewed nutrient concentrations in many small, undisturbed, forested watersheds and found that USEPA guidance criteria were often exceeded, sometimes by as much as 5× the guidance value. Smith et al. (2007) proposed possible nutrient criteria for wadeable streams that were nearly 2× the USEPA guidance values for the same ecoregions. Martinez (2002) cited government standards for Puerto Rico as being 1 mg TP/L, but identified a TP criterion of 0.19 mg/L after using the USEPA percentile method to analyze an independent data set.

Our percentile analysis results and the USEPA's guidance values for aggregate nutrient ecoregions in NY (USEPA 2000a, b, c, 2001) are very similar (our study: 0.04 mg TP/L, 0.7 mg TN/L, 2.4 mg chl *a*/m^3; USEPA guidance: 0.033 mg TP/L, 0.54 mg TN/L, 3.5 mg chl *a*/m^3). This similarity is not surprising given the similarity in method, but it is interesting because the values were derived from 2 independent data sets. Our values are slightly higher than USEPA guidance for both TP and TN, but the degree of similarity among results derived from independent data sets suggests that the results are relatively robust. However, nutrient criteria based on percentiles alone are difficult to connect with water body designated uses and evaluations of use attainment. Thus, additional data sets with information that can be related to classification of water bodies based on use must be included when establishing nutrient criteria.

Use of biological assemblages to set criteria

NYS uses aquatic macroinvertebrates (and recently, periphyton) to assess water quality and to evaluate attainment of aquatic life use of surface waters. Use of these 2 assemblages to establish nutrient criteria bridges the natural environment and the regulatory world. However, not all measures of biotic condition are suited for assessing biological responses to nutrients. In our study, the metrics designed to measure degree of enrichment were most highly correlated with nutrient chemistry. Specifically, NBI-P (macroinvertebrates) and % eutrophic individuals (diatoms) appeared to be the best predictors of

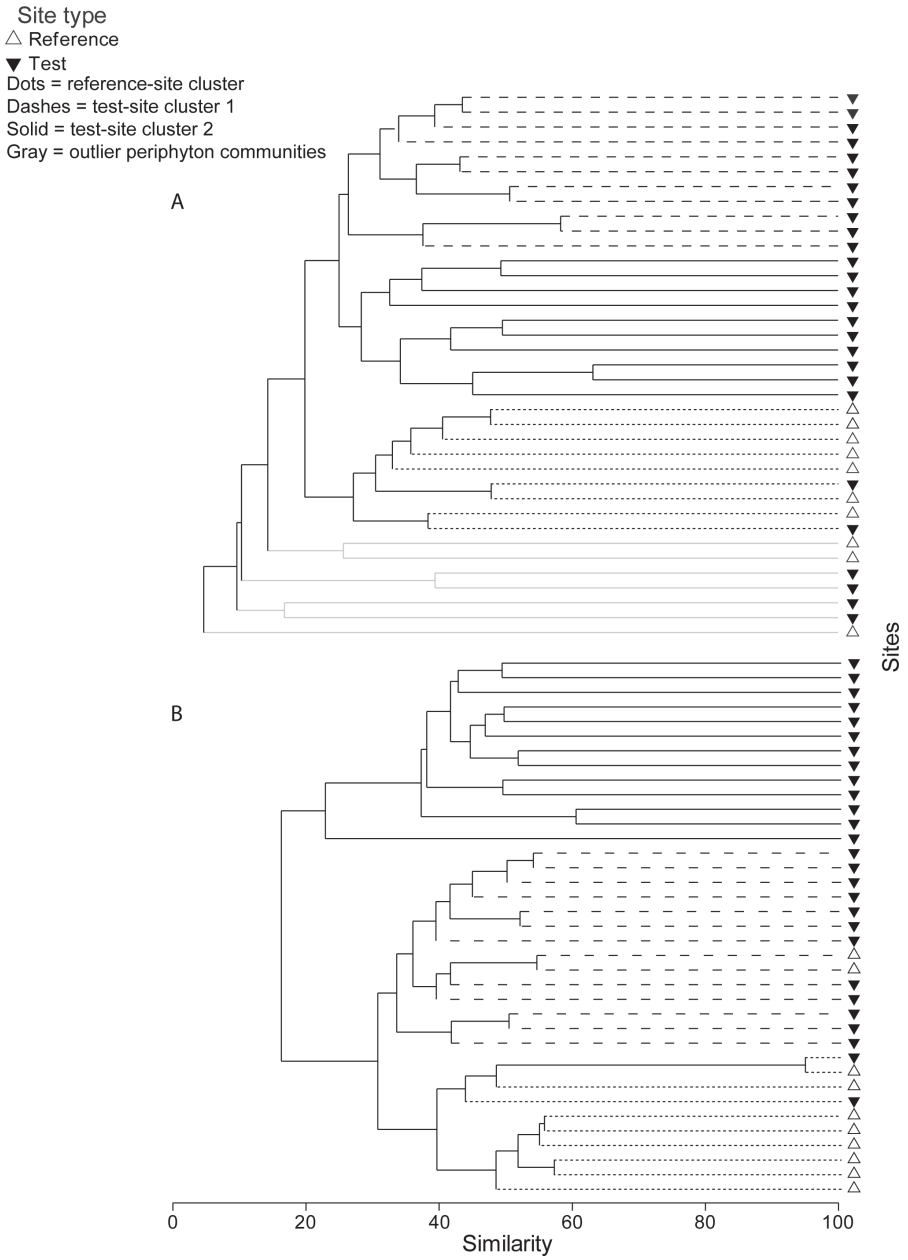

FIG. 5. Dendrograms from Bray–Curtis Similarity Analysis based on log(x)-transformed benthic macroinvertebrate species data (A) and log(x)-transformed periphyton species data (B). Three strong clusters of sites are apparent based on both communities. For both macrophyte and periphyton data, median nutrient concentrations at sites in the reference-site cluster were significantly lower than at sites in test-site cluster 2, and median nutrient concentrations in test-site cluster 1 were intermediate between those at sites in the other 2 clusters (analysis of variance; Figs 6A, B, 7A, B).

increased nutrients in water. Both metrics were significantly correlated with the major nutrient variables, but correlations between % eutrophic individuals and nutrients were stronger than correlations between NBI-P and nutrients (Tables 2, 3). Dodds (2007) suggested that response thresholds related to trophic state boundaries might differ among biological communities in streams and rivers. For example, thresholds were higher for fish and macroinvertebrate communities than for chl *a* (Dodds 2007), probably because nutrients affect primary producers directly and consumers indirectly.

We used biological metrics that were significantly correlated with nutrient concentrations to establish threshold nutrient concentrations that could be applied as criteria. These values were derived from

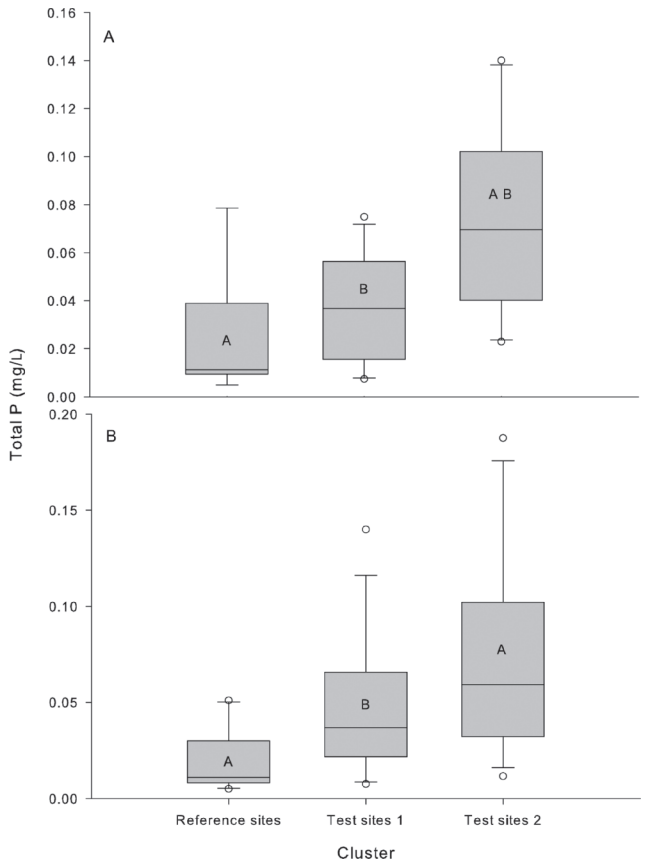

FIG. 6. Box plots of total P concentrations at sites in 3 clusters identified by Bray–Curtis Analysis (Fig. 5A, B) based on benthic macroinvertebrate (A) and periphyton (B) community composition. Lines in boxes are medians, ends of boxes are quartiles, whiskers show 10th and 95th percentiles, and open circles show outliers. Boxes with the same letter are significantly different.

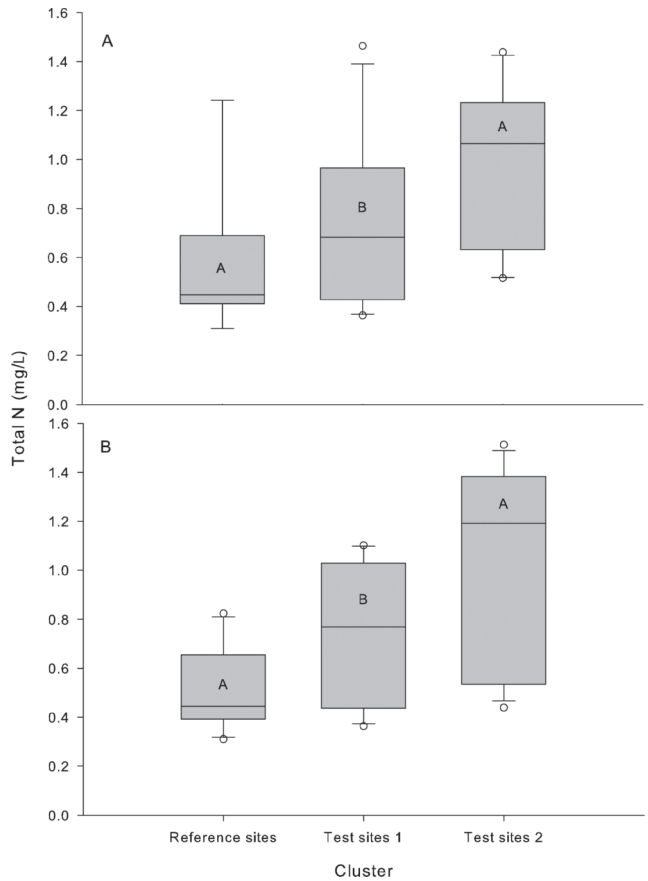

FIG. 7. Box plots of total N concentrations at sites in 3 clusters identified by Bray–Curtis Analysis (Fig. 5A, B) based on benthic macroinvertebrate (A) and periphyton (B) community composition. Lines in boxes are medians, ends of boxes are quartiles, whiskers show 10th and 95th percentiles, and open circles show outliers. Boxes with the same letter are significantly different.

both biological and nutrient data. Thus, they can be more directly tied to aquatic life use attainment and are more meaningful for regulatory purposes than are percentile values. Qian et al. (2003, 2004) used nCPA to identify well-defined TP thresholds above which algal and benthic macroinvertebrate community structure changed. King and Richardson (2003) used the same method to establish TP thresholds above which significant changes occurred in wetland benthic macroinvertebrate communities. King and Richardson (2003) also used a bootstrap resampling technique to generate a frequency distribution of changepoints that provided cumulative probabilities that a changepoint occurred at a given nutrient concentration.

We used nCPA to identify ecological thresholds related to specific nutrient concentrations. However, not all nutrients were equally important, and changepoints differed depending on the biological metric used as the response variable. This variability was directly related to the specific type of stressor the metric was designed to assess and has implications for weighting metrics to define final nutrient criteria. Differences in the mean metric scores between sites with nutrient values greater and less than the threshold are meaningful for water-quality assessment and identifying trophic state. For example, based on NYS BAP scores, sites with TP concentrations below the 50th-percentile changepoint (0.07 mg/L) would be assessed as non- or slightly impacted (Riva-Murray et al. 2002, Smith et al. 2009), whereas sites with TP above the 50th-percentile changepoint would be assessed as moderately or severely impacted, which is within NYS's impairment decision threshold and would trigger remedial action (Smith et al. 2007) and placement on the 303(d) list of impaired waterbodies. Thus, this threshold would not

be protective of aquatic life. Nutrient criteria must be set at values that prevent impairment from eutrophication rather than simply identify it after the damage is done. Based on NBI-P scores, sites with TP above the 50th percentile changepoint (0.011 mg/L) would greatly exceed the eutrophic condition threshold score of 6.0 developed for this index (Smith et al. 2007) (Table 4, Fig. 2A). However, the TP concentration associated with the NBI-P changepoint is significantly smaller than the TP concentration associated with the BAP changepoint, and the NBI-P changepoint would be protective of aquatic life. Therefore, regulatory decisions will have to be made to determine whether criteria are meant to be protective (0.011mg/L TP) or reactive (0.07mg/L) and whether assessment of eutrophication should be placed on the same historical scales of water-quality impairment. Relying more heavily on nutrient-specific metrics, such as the NBI-P, when defining criteria will result in more protective values because they will not result in water-quality impairment. However, they might identify a system that is on the verge of shifting to a higher trophic state and, therefore, is in greater need of protection.

Choice of metrics for setting thresholds

Macroinvertebrate and diatom metrics that were designed specifically to detect nutrient enrichment in streams had threshold responses at nutrient concentrations 7× lower than threshold responses of general water-quality metrics. General water-quality metrics are designed to identify water-quality impairment and are less sensitive than metrics designed to detect nutrient enrichment (Smith et al. 2007). Elevated nutrient concentrations do not necessarily result in impairment in the traditional sense of water quality until at very high concentrations (Sheeder and Evans 2004, Smith et al. 2007). Therefore, nutrient enrichment might create shifts in invertebrate and algal populations at low nutrient concentrations, whereas overall water-quality impairment (as measured by common water-quality metrics) might occur at much higher concentrations. We would argue that nutrient criteria and assessment of nutrient impairment should rely on trophic-state-specific metrics, such as NBI-P or % eutrophic individuals, and that states should use incremental shifts in biological community structure as management thresholds. For example, a TP concentration of 0.07 mg/L probably will result in moderate impact or worse, but changepoints in NBI-P and other metrics indicate that significant shifts in community structure occur at lower TP concentrations. Moreover, preventing impairment is not the only endpoint of many state aquatic life use goals. Preventing shifts in natural ecosystem structure and function of waters also is a management directive. Nutrient criteria that prevent only severe cases of water-quality impairment in which ecosystem structure and function are highly degraded will not protect a state's remaining natural, nondegraded waters. The result of setting criteria based solely on impairment thresholds would be loss of diversity and limited ability to retain natural ecosystem functions, all of which happens at nutrient concentrations much lower than those at which water-quality impairment is measureable.

The BCA detected differences in community composition among sites, and its separation of reference and test sites suggests changes in both benthic macroinvertebrate and diatom community structure. In addition, clusters also reflected nutrient concentrations at sites, and the 3-cluster pattern corresponded to sites with low, medium, and high nutrient concentrations (Fig. 5). This result suggests a possible continuum of trophic states from oligotrophic (low nutrient) to eutrophic (high nutrient), a conclusion supported by similar BCA results in wadeable streams along nutrient gradients (Smith et al. 2007) and general disturbance gradients (Morris et al. 2006, Simon and Morris 2009). The significant difference in nutrient concentrations among clusters of sites reinforces the notion that these sites were separated by more than just chance alone. The median nutrient concentrations of sites within the reference-site (low-nutrient) clusters for both macroinvertebrates (0.011 mg TP/L, 0.45 mg TN/L) and diatoms (0.011mg TP/L, 0.44 mg TN/L) are lower than the results of the percentile analysis (Table 1) and very similar to those of nutrient-specific metrics analyzed with nCPA (Table 4). This result emphasizes the utility and added protection of aquatic life uses that result when more than water chemistry is considered when defining nutrient criteria.

Weight-of-evidence approach

Any of the nutrient values that resulted in measurable responses by the biotic communities could be established as nutrient criteria for large rivers in NYS. However, the values differ and deciding to use only one would neglect the relevant information contained in the others. The weight-of-evidence approach used here combines the results of the various methods used to describe the nutrient–biotic community relationship in large rivers. We used a weighting scale to reduce the subjective nature or the arbitrary assignment of weights. However, BPJ has an important role. We made informed decisions

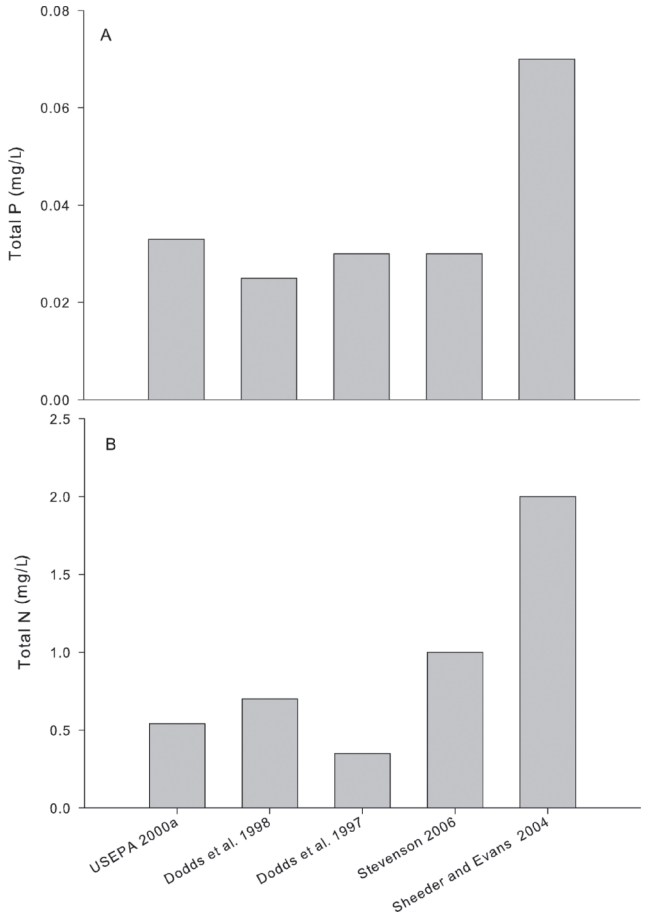

FIG. 8. Proposed nutrient criteria for total P (A) and total N (B) from our study and the literature.

regarding how to weight values based on relevance of the data, the significance of the results, and the perceived protection each value would afford to aquatic life based on the responses of the biotic communities. Future investigators should rely on a similar weighting scale and transparency in their method of assigning various analyses different weights.

The proposed criteria resulting from the weight-of-evidence approach are similar to criteria identified by others to protect aquatic ecosystems from eutrophication (Fig. 8). Dodds et al. (1997) identified 0.03 mg/L TP and 0.35 mg/L TN as necessary in-stream levels to control nuisance benthic chl *a* levels. Dodds et al. (1998) used frequency distributions to establish 0.025 mg TP/L and 0.7 mg TN/L as the boundary between oligotrophic and mesotrophic conditions. USEPA guidance values for the nutrient ecoregions we evaluated suggest criteria of 0.033 mg TP/L and 0.54 mg TN/L (USEPA 2000a). In addition, ranges of natural background nutrient concentrations for TP and TN were established for rivers and headwaters of the US by Smith et al. (2003). The upper limits of these ranges (0.03 mg TP/L and 0.7 mg TN/L) for rivers in NYS (Smith et al. 2003) were almost identical to those proposed here. Some researchers have proposed slightly higher but similar values to ours (Havens 2003, Sheeder and Evans 2004, Stevenson et al. 2006). Stevenson et al. (2006) compared the effects of nutrients on algal biomass in streams and found that most responses occurred between 0.01 to 0.03 mg TP/L and 0.4 to 1.0 mg TN/L. The values we proposed here fall within these ranges. Sheeder and Evans (2004) identified thresholds of impairment in a Pennsylvania watershed that were 0.07 mg TP/L and 2.0 mg TN/L, both higher than the criteria proposed here.

Previous work in NYS on wadeable stream nutrient criteria proposed values of 0.065 mg TP/L and 0.98 mg NO_3-N/L (Smith et al. 2007). These concentrations are high relative to the criteria proposed here for large rivers (Fig. 8). The disparity between values probably stems from using different evaluation methods rather than from differences in type of water body (i.e., large river vs wadeable stream). Smith et al. (2007) evaluated benthic macroinvertebrate responses to nutrients without considering diatom communities and did not use nCPA to evaluate responses of different metrics. Instead nutrient levels were defined as the mesotrophic–eutrophic boundary, which was the concentration at which most impairment occurred, based on BAP scores. However, setting the criteria at these concentrations and impairment levels would not be suitably protective. BAP scores are not as well connected to measuring the response to nutrients as other metrics, and the weight-of-evidence approach allows inclusion of other metrics and biotic assemblages. The values proposed here for large rivers are more similar to those of the oligotrophic–mesotrophic boundary in this previous stream investigation (0.0175 mg TP/L, 0.24 mg NO_3-N/L) (Smith et al. 2007).

Establishing protective, meaningful nutrient criteria is imperative to prevent further eutrophication of surface waters. The use of a weight-of-evidence approach allows inclusion of values obtained by many methods of establishing protective nutrient concentrations. The techniques we used are adaptable and yield easily perceived criteria that can be applied readily to other nutrient-criteria development efforts throughout the US and globally. The available information on large-river biotic communities is limited everywhere. Thus, our work is a substantial starting point for proposing nutrient criteria in large rivers and is a valuable source of information for further study on the relationships between nutrients

and biological assemblages, such as macroinvertebrates and diatoms. The proposed nutrient criteria will protect large river biotic communities and help maintain current trophic condition of rivers in NYS.

Acknowledgements

We thank the dedicated staff of the NYSDEC Stream Biomonitoring Unit for their assistance in sample collection and sample processing, especially Robert Bode and Margaret Novak for their undoubted support; the professional staff of Ecoanalysts, Inc. (Moscow, Idaho), especially Gary Lester, Julia Kay Eichman, and Dave Langill for processing diatom samples; Susy King and the staff of New England Interstate Water Pollution Control Commission for their unrelenting patience during a long process while tracking the funds that made the project possible; the USEPA Region 1 staff, especially Izabela Wojtenko for her assistance in guiding the direction of this project from before its inception; and Mike Paul and Karen Riva-Murray for their insightful comments in early reviews of this manuscript. In addition, we would like to extend special thanks to Ryan King for his assistance in analyzing our data set using nCPA and for permission to use his format for presentation of changepoint results.

Literature Cited

BRAY, J. R., AND J. T. CURTIS. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs 27:325–349.

DAVIS, J. F. 2002. Factors affecting photosynthetic rates of periphyton in shallow streams of southeastern Pennsylvania. Water Environment Research 74:370–376.

DODDS, W. K. 2007. Trophic state, eutrophication and nutrient criteria in streams. Trends in Ecology and Evolution 22:669–676.

DODDS, W. K., J. R. JONES, AND E. B. WELCH. 1998. Suggested classification of stream trophic state: distributions of temperate stream types by chlorophyll, total nitrogen, and phosphorus. Water Research 32:1455–1462.

DODDS, W. K., V. H. SMITH, AND B. ZANDER. 1997. Developing nutrient targets to control benthic chlorophyll levels in streams: a case study of the Clark Fork River. Water Research 31:1738–1750.

DODDS, W. K., AND E. B. WELCH. 2000. Establishing nutrient criteria in streams. Journal of the North American Benthological Society 19:186–196.

FLOTEMERSCH, J. E., J. B. STRIBLING, AND M. J. PAUL. 2006. Concepts and approaches for the bioassessment of non-wadeable streams and rivers. EPA-600-R-06-127. Office of Science and Technology, Office of Research and Development, US Environmental Protection Agency, Washington, DC.

FORE, L. S., AND C. GRAFE. 2002. Using diatoms to assess the biological condition of large rivers in Idaho (U.S.A.). Freshwater Biology 47:2015–2037.

HAVENS, K. E. 2003. Phosphorus–algal bloom relationships in large lakes of south Florida: implications for establishing nutrient criteria. Lake and Reservoir Management 19:222–228.

HESTER, F. E., AND J. S. DENDY. 1962. A multiple-plate sampler for aquatic macroinvertebrates. Transactions of the American Fisheries Society 91:420–421.

HILSENHOFF, W. L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20:31–40.

HUGHES, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31–48 *in* W. S. Davis and T. P. Simon (editors). Biological assessment and criteria: tools for water resource planning. CRC Press, Boca Raton, Florida.

ICE, G., AND D. BINKLEY. 2003. Forest streamwater concentrations of nitrogen and phosphorus: a comparison with EPA's proposed water quality criteria. Journal of Forestry 101:21–28.

KING, R. S., M. E. BAKER, D. F. WHIGHAM, D. E. WELLER, T. E. JORDAN, P. F. KAZYAK, AND M. K. HURD. 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15:137–153.

KING, R. S., W. V. DELUCA, D. F. WHIGHAM, AND P. P. MARRA. 2007. Threshold effects of coastal urbanization on *Phragmites australis* (common reed) abundance and foliar nitrogen in Chesapeake Bay. Estuaries and Coasts 30:469–481.

KING, R. S., AND C. J. RICHARDSON. 2003. Integrating bioassessment and ecological risk assessment: an approach to developing numerical water-quality criteria. Environmental Management 31:795–809.

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

LENAT, D. R. 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. Journal of the North American Benthological Society 7:222–233.

MAGURRAN, A. E. 2004. Measuring biological diversity. African Journal of Aquatic Science 29:285–286.

MARTINEZ, G. A. 2002. Development of numeric nutrient criteria for lakes in Puerto Rico. Journal of Agriculture–University Of Puerto Rico 86:139–144.

MAXTED, J. R., M. T. BARBOUR, J. GERRITSEN, V. PORETTI, N. PRIMROSE, A. SILVIA, D. PENROSE, AND R. RENFROW. 2000. Assessment framework for mid-Atlantic Coastal Plain streams using benthic macroinvertebrates. Journal of the North American Benthological Society 19:128–144.

MILTNER, R. J., AND E. T. RANKIN. 1998. Primary nutrients and the biotic integrity of rivers and streams. Freshwater Biology 40:145–158.

MORRIS, C. C., T. P. SIMON, AND S. A. NEWHOUSE. 2006. A local-scale in situ approach for stressor identification of biologically impaired aquatic systems. Archives of

Environmental Contamination and Toxicology 50: 325–334.

PALMSTROM, N. 2002. Development of regional nutrient criteria and implications for the regulated community. Proceedings of the Water Environment Federation 2002: 532–540.

PARSONS, M., AND R. H. NORRIS. 1996. The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. Freshwater Biology 36: 419–434.

PAUL, J. F., AND M. E. MCDONALD. 2005. Development of empirical, geographically specific water quality criteria: a conditional probability analysis approach. Journal of the American Water Resources Association 1211–1223.

QIAN, S. S., R. S. KING, AND C. J. RICHARDSON. 2003. Two statistical methods for the detection of environmental thresholds. Ecological Modelling 166:87–97.

QIAN, S. S., Y. PAN, AND R. S. KING. 2004. Soil total phosphorus threshold in the everglades: a Bayesian changepoint analysis for multinomial response data. Ecological Indicators 4:29–37.

RECKHOW, K. H., G. B. ARHONDITSIS, M. A. KENNEY, L. HAUSER, J. TRIBO, C. WU, K. J. ELCOCK, L. J. STEINBERG, C. A. STOW, AND S. J. MCBRIDE. 2005. A predictive approach to nutrient criteria. Environmental Science and Technology 39:2913–2919.

REYNOLDSON, T. B., R. H. NORRIS, V. H. RESH, K. E. DAY, AND D. M. ROSENBERG. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society 16:833–852.

RIVA-MURRAY, K., R. W. BODE, P. J. PHILLIPS, AND G. L. WALL. 2002. Impact source determination with biomonitoring data in New York State: concordance with environmental data. Northeastern Naturalist 9:127–162.

ROHM, C. M., J. M. OMERNIK, A. J. WOODS, AND J. L. STODDARD. 2002. Regional characteristics of nutrient concentrations in streams and their application to nutrient criteria development 1. Journal of the American Water Resources Association 38:213–239.

SHEEDER, S. A., AND B. M. EVANS. 2004. Estimating nutrient and sediment threshold criteria for biological impairment in Pennsylvania watersheds. Journal of the American Water Resources Association 40:881–888.

SIMON, T. P., AND C. C. MORRIS. 2009. Biological response signature of oil brine threats, sediment contaminants, and crayfish assemblages in an Indiana watershed, USA. Archives of Environmental Contamination and Toxicology 56:96–110.

SMITH, A. J., R. W. BODE, AND G. S. KLEPPEL. 2007. A nutrient biotic index (NBI) for use with benthic macroinvertebrate communities. Ecological Indicators 7:371–386.

SMITH, A. J., D. L. HEITZMAN, AND B. T. DUFFY. 2009. Standard operating procedure: biological monitoring of surface waters in New York State. Stream Biomonitoring Unit, New York State Department of Environmental Conservation, Troy, New York. (Available from: New York State Department of Environmental Conservation, 425 Jordan Road, Troy, New York 12180 USA.)

SMITH, R. A., R. B. ALEXANDER, AND G. E. SCHWARZ. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous United States. Environmental Science and Technology 37:3039–3047.

SNELDER, T. H., B. J. F. BIGGS, AND M. A. WEATHERHEAD. 2004. Nutrient concentration criteria and characterization of patterns in trophic state for rivers in heterogeneous landscapes. Journal of the American Water Resources Association 40:1–13.

STEVENSON, R. J., S. T. RIER, C. M. RISENG, R. E. SCHULTZ, AND M. J. WILEY. 2006. Comparing effects of nutrients on algal biomass in streams in two regions with different disturbance regimes and with applications for developing nutrient criteria. Hydrobiologia 561:149–165.

TETRA TECH, INC. 2008. Draft nutrient criteria to protect aquatic life uses in Mississippi non-tidal streams and rivers. Mississippi Department of Environmental Quality, Office of Pollution Control, Jackson, Mississippi. (Available from: Tetra Tech, Inc. Center for Ecological Sciences, 400 Red Brook Blvd., Suite 200, Owings Mills, Maryland 21117-5159 USA).

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 1993. Methods for the determination of inorganic substances in environmental samples. EPA 600-R-93-100. Office of Research and Development, US Environmental Protection Agency, Cincinnati, Ohio.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 1998. National strategy for the development of regional nutrient criteria. EPA 822-R-98-002. Office of Water, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000a. Ambient water quality criteria recommendations, information supporting the development of state and tribal nutrient criteria, rivers and streams in nutrient ecoregion VII. 93. EPA 822-B-00-018. Office of Science and Technology, Office of Water, Health and Ecological Criteria Division, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000b. Ambient water quality criteria recommendations, information supporting the development of state and tribal nutrient criteria, rivers and streams in nutrient ecoregion XI. EPA 822-B-00-020. Office of Science and Technology, Office of Water, Health and Ecological Criteria Division, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000c. Ambient water quality criteria recommendations, information supporting the development of state and tribal nutrient criteria, rivers and streams in nutrient ecoregion XIV. EPA 822-B-00-022. Office of Science and Technology, Office of Water, Health and Ecological Criteria Division, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000d. Atlas of America's polluted waters. EPA 840-B00-002.

Office of Science and Technology, Office of Water, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000e. Nutrient criteria technical guidance manual, rivers and streams. EPA-822-B-00-002. Office of Science and Technology, Office of Water, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2001. Ambient water quality criteria recommendations, information supporting the development of state and tribal nutrient criteria, rivers and streams in nutrient ecoregion VIII. EPA 822-B-01-015. Office of Science and Technology, Office of Water, Health and Ecological Criteria Division, US Environmental Protection Agency, Washington, DC.

VAN DAM, H., A. MERTENS, AND J. SINKELDAM. 1994. A coded checklist and ecological indicator values of freshwater diatoms from The Netherlands. Aquatic Ecology 28:117–133.

WANG, L., D. M. ROBERTSON, AND P. J. GARRISON. 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: implication to nutrient criteria development. Environmental Management 39:194–212.

WILHELM, J. G. O., J. D. ALLAN, K. J. WESSELL, R. W. MERRITT, AND K. W. CUMMINS. 2005. Habitat assessment of nonwadeable rivers in Michigan. Environmental Management 36:592–609.

Received: 15 June 2009
Accepted: 21 April 2010