J. N. Am. Benthol. Soc., 2008, 27(3):783–799

DOI: 10.1899/07-077.1

Algae–P relationships, thresholds, and frequency distributions guide nutrient criterion development

---

R. Jan Stevenson[1]
Department of Zoology, Michigan State University, East Lansing, Michigan 48824 USA

Brian H. Hill[2]
National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division, US Environmental Protection Agency, Duluth, Minnesota 55804-2595 USA

Alan T. Herlihy[3]
Department of Fisheries and Wildlife, Oregon State University, c/o US Environmental Protection Agency, Corvallis, Oregon 97333 USA

Lester L. Yuan[4] AND **Susan B. Norton**[5]
National Center for Environmental Assessment, Office of Research and Development, US Environmental Protection Agency, Washington, DC 20460 USA

Abstract. We used complementary information collected using different conceptual approaches to develop recommendations for a stream nutrient criterion based on responses of algal assemblages to anthropogenic P enrichment. Benthic algal attributes, water chemistry, physical habitat, and human activities in watersheds were measured in streams of the Mid-Atlantic Highlands region as part of the Environmental Monitoring and Assessment Program of the US Environmental Protection Agency. Diatom species composition differed greatly between low- and high-pH reference streams; therefore, analyses for criterion development were limited to a subset of 149 well-buffered streams to control for natural variability among streams caused by pH. Regression models showed that TP concentrations were ~10 μg/L in streams with low levels of human activities in watersheds and that TP increased with % agriculture and urban land uses in watersheds. The 75th percentile at reference sites was 12 μg TP/L. Chlorophyll *a* and ash-free dry mass increased and acid and alkaline phosphatase activities decreased with increasing TP concentration. The number of diatom taxa, evenness, proportion of expected native taxa, and number of high-P taxa increased with TP concentration in streams. In contrast, the number of low-P native taxa and % low-P individuals decreased with increasing TP. Lowess regression and regression tree analysis indicated nonlinear relationships for many diversity indices and attributes of taxonomic composition with respect to TP. Thresholds in these responses occurred between 10 and 20 μg/L and helped justify recommending a P criterion between 10 and 12 μg TP/L to protect high-quality biological conditions in streams of the Mid-Atlantic Highlands.

Key words: algae, biomass, diatoms, frequency distributions, Mid-Atlantic Highlands, nutrients, nutrient criteria, phosphorus, species composition, streams, stressor–response relationship, threshold.

Nutrient enrichment by humans is recognized as a leading cause of degradation of the ecological integrity of streams in the US (Vitousek et al. 1997, Carpenter et al. 1998, Smith et al. 1999, USEPA 2000). Establishment of nutrient criteria has become the focus of a national effort to regulate nutrient content in all water body types in all regions of the US (USEPA 1998). This national effort has resulted in the production of new information, syntheses of old information, and devel-

---

[1] E-mail addresses: rjstev@msu.edu
[2] hill.brian@epamail.epa.gov
[3] herlihy.alan@epa.gov
[4] yuan.lester@epamail.epa.gov
[5] norton.susan@epamail.epa.gov

opment of new concepts that link the basic and applied sciences. Questions arose about the use of different approaches, such as frequency distributions and stressor–response relationships, recommended by the USEPA (2000) for development of nutrient criteria for streams. Our paper addresses use of those approaches.

Nutrient enrichment of streams caused by human activities affects algal biomass and species composition. Accrual of benthic algal biomass is limited by low P and N concentrations in many streams (Francoeur 2001), and algal biomass is positively related to nutrient concentrations in streams (Biggs 2000, Dodds et al. 2002, Stevenson et al. 2006). Nutrient concentrations also affect relative abundances of diatom species in streams (Kelly and Whitton 1995, Pan et al. 1999, Manoylov and Stevenson 2006).

We often expect algal species richness and evenness to decrease in streams with increasing pollution, but nutrients also are a resource stressor. Increasing resource availability from low to medium levels can increase diversity because low nutrient concentrations might constrain algal taxonomic composition to those species adapted to low nutrient concentrations (Manoylov and Stevenson 2006). Thus, release from nutrient limitation might allow increases in species richness in physically disturbed systems such as streams (Biggs and Smith 2002, Huston 2004). Despite this potential increase in overall species richness, addition of nutrients could reduce relative abundance and richness of native species by altering competitive hierarchies or habitat structure (Huston 1994). These relationships between biodiversity and nutrients have not been evaluated thoroughly for algae in streams.

Several approaches to establishing environmental criteria have been recommended, and these approaches provide complementary information (USEPA 2000, Stevenson et al. 2004). Regression models relating nutrient concentration and land use in watersheds can be used to predict natural or best available nutrient concentrations in streams (Dodds and Oakes 2004). Documenting the relationships between nutrients and human activities is important for distinguishing natural and anthropogenic sources of variability in nutrients among streams. Frequency distributions of nutrient concentrations at reference sites, or at sites meeting other management criteria, such as high invertebrate biodiversity, can be used to characterize the variability in nutrients at sites that do meet management expectations. Stressor–response relationships characterize the progressive degradation in ecological systems with increasing pollution (stressors, e.g., nutrients) and help delineate the pollution concentrations that would support valued ecological attributes.

Relationships between nutrients and valued ecological attributes are important for documenting effects on ecosystems that generate support from stakeholders for management actions. For example, water clarity and algal blooms caused by nutrients are important concerns for stakeholders. Thus, the stressor–response approach can be used to establish effects-based criteria by quantitatively relating nutrient concentrations to valued ecological attributes (e.g., water clarity) and determining the specific benefits (e.g., fewer algal blooms) of protecting or restoring ecosystems at specific levels of nutrient concentrations. Criteria are established with this *stressor–response* approach at levels of stressors that provide the desired level of protection for the valued ecological attribute. These stressor–response relationships might be developed based on laboratory bioassays or field surveys and are the foundation of risk assessment (Suter 1993, USEPA 1998). Thresholds (i.e., sudden changes in attributes along the stressor gradients) can be particularly valuable for justifying criteria at specific stressor levels (Muradian 2001). In the development of P criteria for the Everglades, threshold effects on calcareous algal mats and species composition of biota were important for justifying a specific P concentration that would protect biological integrity (e.g., Stevenson et al. 2002, King and Richardson 2003). Some evidence for quantitative relationships between nutrients and algal biomass has been reported for streams, but few studies have evaluated thresholds in relationships between nutrients and algal attributes.

The objectives of our research were to relate changes in stream benthic algae to anthropogenic increases in total P (TP) concentrations and to develop recommendations for TP criteria. We chose TP rather than total N (TN) or both TP and TN to limit complexity of the results and because experiments, nutrient ratios, and correlations among attributes indicate that P commonly regulates algal responses in streams of the study region (e.g., Pan et al. 1999, Francoeur 2001). First, we establish relationships between TP concentrations in streams and human activity in watersheds to predict natural TP concentrations in streams (predictive models). We also use frequency distributions of TP concentrations at sites with low levels of human activities to characterize natural TP concentrations. Then we relate TP to changes in algal biomass, enzyme activity, diversity, and taxonomic composition in these streams to confirm that TP is correlated with changes in valued ecological attributes, to characterize magnitudes of effects, and to identify thresholds in algae–TP relationships. We identify candidate concentrations for

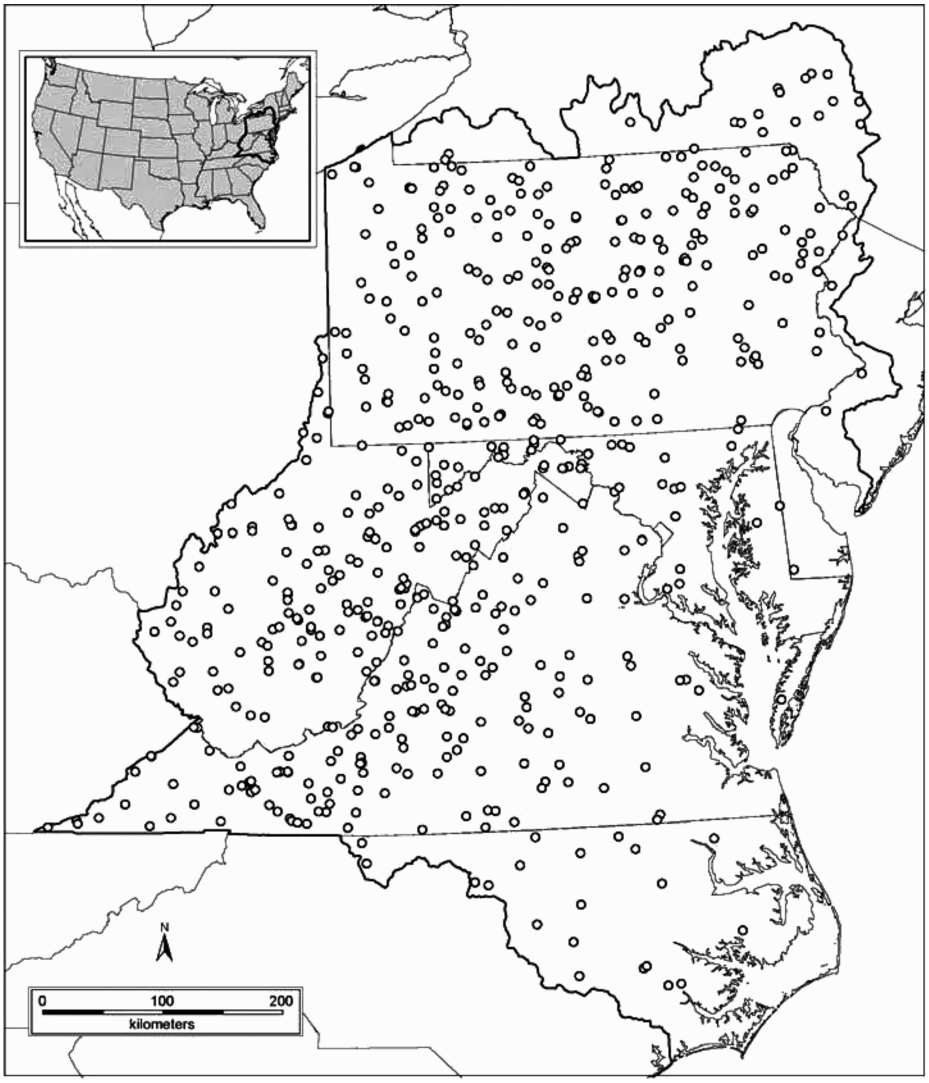

FIG. 1. Sites in the Mid-Atlantic Highlands region of the US, including Pennsylvania, West Virginia, Maryland, Virginia, and parts of North Carolina and New York, where periphyton samples were collected during 1993 to 1998.

TP criteria (benchmarks) using predictive models of natural conditions, frequency distributions, stressor–response relationships, and thresholds. Last, we compare these benchmarks and discuss transferability among regions.

Methods

Sample collection and analysis

From 1993 to 1998, 1185 periphyton samples were collected in 607 streams throughout the Mid-Atlantic Highlands (MAH) region as part of the Environmental Monitoring and Assessment Program (EMAP) of the US Environmental Protection Agency (EPA) (Fig. 1). This region covers Maryland, Pennsylvania, West Virginia, Virginia, Delaware, and parts of New York and North Carolina. It includes 7 ecoregions (Omernik 1987) and 5 major watersheds (Alleghany, Kanawha, Monongahela, Potomac, and Susquehanna). Periphyton was collected by scraping and brushing organic material from a known area of natural substrate at a random location on each of 9 transects in a stream reach (Pan et al. 1996, Hill et al. 2000; see Lazorchak et al. 1998 for details of all methods). Samples from riffles and pools were separated into 2 composite samples from a stream if both habitats existed in the reach. The volume of each sample was measured in the field, and the sample was then split into 4 subsamples for assays of chlorophyll *a* (chl *a*), ash-free dry mass (AFDM), diatom taxonomic composition, and acid and alkaline phosphatase activity. Chl *a* was measured with a spectrophotometer using standard methods (APHA

1992) including acidification to determine the phaeophytin degradation portion of uncorrected chl *a* estimates. AFDM was assayed by drying and combusting samples according to standard methods (APHA 1992). Diatom taxonomic composition was determined by identifying and counting ≥500 diatom valves at 1000× after cleaning and mounting diatoms on slides in a high refractive index mounting medium (HYRAX® or NAPHRAX®; The Biology Shop, Hazelbrook, New South Wales, Australia; http://www.pnc.com.au/~biology). Acid and alkaline phosphatases in periphyton subsamples were assayed with standard methods and using nitrophenyl phosphate as a substrate (Sayler et al. 1979).

Water-chemistry, physical-habitat, and riparian characteristics in watersheds also were assessed at most sites. A list of sample preparation and preservation protocols for water chemistry can be found in Pan et al. (1996). Water-chemistry characteristics, such as nutrients, Cl, pH, SO_4, and many other variables, were analyzed with procedures outlined in USEPA (1987). TP and TN were analyzed spectrophotometrically after persulfate digestion. Soluble reactive P was not measured, but NO_3 and NH_4 were. Physical-habitat characteristics measured in streams included thalweg depth, mean wetted width, mean reach cross section, slope, residual pool area, substrate size, and embeddedness (Kaufmann et al. 1999). Land use and land cover were characterized for topographically delineated watersheds with US Geological Survey landuse and land-cover digital maps and Thematic Mapper satellite imagery (Herlihy et al. 1998, Hill et al. 2000).

Data analysis

We downloaded the data from the EMAP web site (www.epa.gov/emap). Diatom data consisted of cell densities. We calculated diatom relative abundances (% of diatoms counted) by dividing cell densities of each species by sums of cell densities for all diatom taxa in the same sample and multiplying by 100. We restricted analyses of relationships among nutrients, diatoms, and human activities to samples collected from riffles during 1993 to 1996, when most samples came from 1st- to 3rd-order streams in the MAH. The 1997 to 1998 samples came from both wadeable and nonwadeable streams in the Coastal Plain and Piedmont ecoregions of the Mid-Atlantic states. Some streams in the database had been sampled more than once during the year and in more than one year. In those cases, we used the first sample from the first year in which the stream was sampled. Data from the streams sampled during 1997 to 1998 were used to develop diatom autecological information (see *Effects-based TP benchmarks* below).

Stream classification.—The 1st step in data analysis was to classify MAH streams based on similarity in diatom taxonomic composition among reference sites so that we could limit analyses to streams in which we expected similar natural (or reference) conditions and similar responses to watershed disturbance. We chose reference sites initially for consistency with other analyses of MAH streams based on EMAP data. We used selection criteria that were assumed to eliminate sites with evidence of stressors from mining, agriculture, and urban development (criteria and assumptions are described in Waite et al. 2000). These criteria were >50 μeq/L acid neutralizing capacity (ANC), <100 μeq Cl/L, <20 μg TP/L, <750 μg TN/L, and <400 μeq SO_4/L. We included species in the cluster analysis if their relative abundances averaged >1% at the reference sites where they were observed. Two-way indicator species analysis (TWINSPAN, www.canodraw.com/wintwins.htm; Hill et al. 1975) was used to group sites with similar diatom assemblages because this analysis provided groups of sites that could be related to environmental factors in past studies (e.g., Pan et al. 2000).

We used linear discriminant analyses and descriptive statistics (SYSTAT 10.0; Systat Software, Chicago, Illinois) to determine the natural ecological variables that could best predict membership of sites in different TWINSPAN groups. The 1st TWINSPAN analysis indicated that the person counting samples had a significant effect on membership of sites in TWINSPAN groups. Therefore, we reviewed the list of diatom taxa to determine the probable differences in how analysts had identified taxa. We combined groups of taxa with similar morphology, such as varieties within the same species and groups of small taxa. After lumping taxa according to the preceding criteria, we analyzed the data again; then environmental factors that often affect diatom species composition became most important for distinguishing groups of reference sites. Subsequent discriminant analyses with forward and backward selection identified environmental variables that best distinguished membership of sites in TWINSPAN groups. We chose a group of streams with the largest number of sites sampled during the 1993 to 1996 field seasons for all analyses of relationships among nutrients, algae, and human activities. We chose this level of classification because finer levels of classification (more small groups) did not improve relationships between metrics and TP concentration.

TP benchmarks for best available conditions.—*Benchmarks* is a term used in ecotoxicological literature to refer to specific levels of variables that can be justified, for one reason or another, as candidates for criteria (Setzer and Kimmel 2003). The actual criterion is then selected from the benchmarks. We determined benchmarks related to best available TP conditions using 2 different methods. First, we used linear regression to relate TP to the percentages of watersheds altered by human activities (% WA) (SYSTAT). We calculated % WA as the sum of % agricultural, urban, and mining land uses. We used the resulting model to document that human activities affected TP concentration and to predict TP concentration when human activities were lowest in the region, i.e., when % WA was 0. Preliminary exploration of the data indicated that no more variation in TP could be explained by separating % agricultural, urban, and mining land uses in the regression model than by using % WA. We set the 1st benchmark for TP criteria at the 75% confidence interval (CI) for predicted TP concentration when % WA was 0. We chose the 75% CI rather than the 90th and 95th to balance Type I and Type II errors in the same way that selection of the 75th percentile of frequency distributions balances these errors (see below).

We used the 75th percentile of frequency distributions of TP concentrations for reference sites as the 2nd benchmark for TP criteria because of precedent and USEPA recommendations (Dodds et al. 1998, USEPA 2000). Different definitions of reference condition, such as best professional judgment for conditions supporting high biological diversity of aquatic invertebrates or sites that actually support biological conditions that are deemed acceptable, have been used. After the preliminary stream classification step (see *Stream classification* above), we modified our definition of reference condition for nutrient criteria development. We used 0% WA as a conservative starting point for selecting reference sites because these conditions existed in the region. The number of streams with 0% WA was very low, so we repeated evaluations of TP frequency distributions with the maximum % WA set at successively higher and higher levels (% WA $\leq$0, 1, 2, 3, 4, 5, 6, 8, 10, 15, 20, 30, 40%). We then plotted the median and 75th percentile TP concentrations of the TP frequency distributions as a function % WA in the 13 sets of streams. We used the largest set of streams in which the median and 75th percentile of TP concentrations did not increase above natural (i.e., concentrations at 0% WA) to define the maximum % WA for reference sites and to determine the 75th percentile of the TP concentrations in reference condition. We used the same statistical analysis for TN and Cl (Cl is an indicator of human activities in watersheds) to evaluate corresponding changes in other chemical conditions as human activities increased in watersheds.

Effects-based TP benchmarks.—We calculated stressor–response relationships with 11 algal attributes (described below) to determine specific effects of specific TP concentrations. We selected the 11 algal attributes because they were indicators related to valued ecological attributes, such as intrinsic values, nuisance problems, or periphyton function.

We included 2 biomass indicators, chl *a* and ash-free dry mass (AFDM) per unit area, as algal attributes because they are 2 common measures of benthic algal biomass in streams, and they were provided in the EPA database. Alkaline and acid phosphatase activities, indicators of P demand by microbes in aquatic ecosystems, also were provided in the EPA database. We calculated 4 diversity indicators with species data: the number of taxa observed in the 500-valve count, evenness (Hurlbert 1971), proportion of expected native taxa (see definition below), and the proportion of expected native taxa able to live in low P concentrations (low-P native taxa). Last, we calculated 3 measures of taxonomic composition: Bray–Curtis (BC) similarity of species composition to reference condition, proportion of taxa able to live in low P concentrations (low-P taxa), and proportion of taxa requiring high P (high-P taxa).

We defined native taxa as the taxa occurring at reference sites. Many of these taxa are commonly found at sites around the US where % WA is low; thus, they probably existed in North American streams prior to major alterations of the landscape by humans. We identified native taxa originally from a randomly selected subset consisting of ½ of the reference sites (subset 1). We measured the proportion of native taxa at a test (nonreference) site as the ratio of the observed number of subset-1 native taxa at the test site to the expected number of native taxa at a subset-1 reference site. The observed number of subset-1 native taxa at a test site was the number of native taxa at the site that had been observed in subset-1 reference sites. The expected number of native taxa at a test site was the average number of native taxa at subset-1 reference sites. We used the native taxa observed in the other ½ of the reference sites (subset 2) to measure the proportion of native taxa at a subset-1 reference site as the ratio of the observed number of subset-2 native taxa at the subset-1 reference site to the average number of native taxa at subset-2 reference sites. The proportion of native taxa varies from 0 to >1.0. Values are ~1.0 at high-quality sites, <1.0 when native taxa have been lost, and >1.0 when more native species are present than would be expected at reference sites (thus, more native taxa than naturally would occur at a site). This metric is the same as the ratio of observed to expected species number used in River InVertebrate

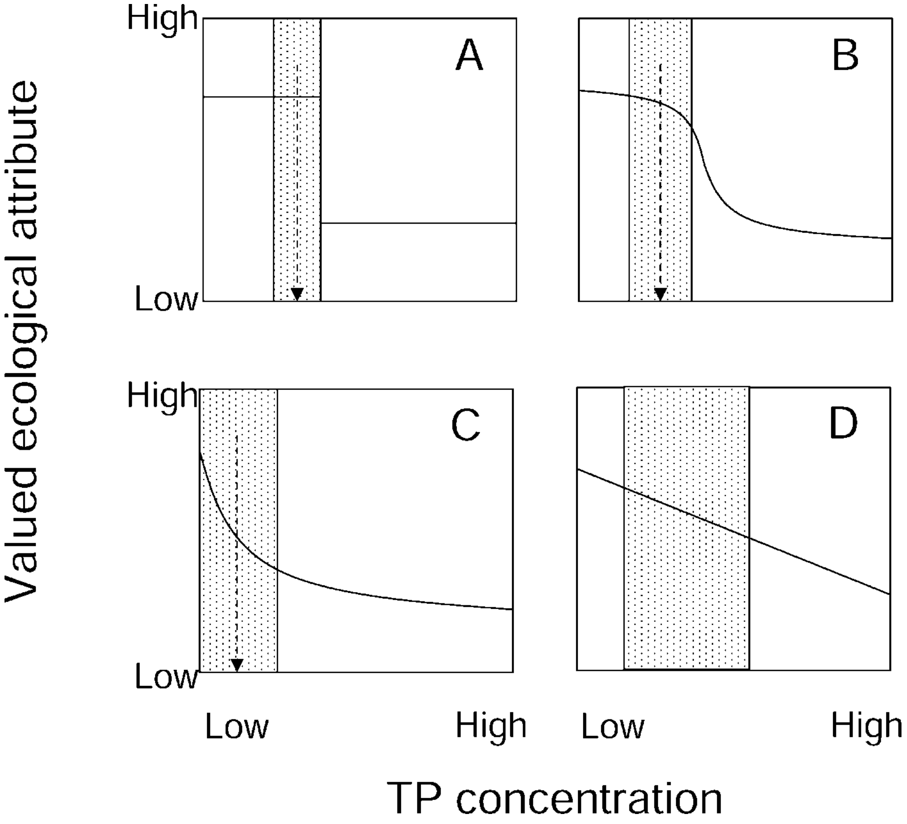

FIG. 2. Approaches to development of stressor criteria when potential responses of valued ecological attributes to stressors (e.g., total P [TP]) are nonlinear with assimilative capacity for increases at low levels the stressor (A, B), nonlinear with strong sensitivity to changes at low levels of the stressor (C), and linear (D). A stressor criterion is established at a level of a stressor that protects the valued ecological attribute. Arrows indicate TP criteria justified on the basis of the form of the stressor–response relationship. Shaded areas indicate the range of TP criteria that could be acceptable. Acceptable ranges vary as a function of the linearity of the stressor–response relationship and the type of nonlinear relationship.

Prediction And Classification System (RIVPACS) models and invertebrate assessment (Wright et al. 1993, Hawkins et al. 2000), but in this case, we studied only one class of streams. RIVPACS models are usually used for many classes of streams simultaneously.

We calculated similarity of taxonomic composition to reference condition with a BC similarity index (i.e., 1 – BC dissimilarity; Legendre and Legendre 1998). For each site, we calculated BC similarity between the site and each reference site and used the mean of these BC similarities to indicate the similarity of the site to reference condition. BC similarity to reference condition ranges from 0.0 (no individuals of the same species are found in the compared assemblages) to 1.0 (relative abundances of all taxa at 2 sites are the same).

We used relative abundances of diatoms and TP concentrations in streams during the 1997 to 1998 EMAP sampling seasons to determine TP autecologies for the diatoms in the 1993 to 1996 data set to maintain independence in analyses. We used weighted average regression (CALIBRATE, version 0.6®; Juggins and ter Braak 1992) to determine species TP optima. We transformed TP optima into 6 categorical values so that these data could be combined with another database of categorical autecological characterizations (van Dam et al. 1994). Values for TP optima ranged from 1 for species with the lowest TP optima to 6 for species with the highest TP optima. We defined low- and high-P taxa as taxa with TP autecologies in the 2 lowest or 2 highest TP optimum categories, respectively. We calculated % low- and % high-P individuals by adding the relative abundances of all low- or high-P taxa, respectively, in a sample. We also calculated % low-P native taxa by adding the relative abundances of those low-P taxa that had been characterized as native in a sample.

We used linear regression, lowess regression, and regression tree analysis to relate TP concentration to algal attributes to determine effects-based TP benchmarks. We used the 3 regression methods as complementary approaches (Fig. 2A–D). Regression tree and lowess techniques order observations along the stressor gradient. Both methods provide analyses of the stressor–response pattern without the a priori assumption of a specific nonlinear pattern.

Regression tree analysis can identify precise change-points (thresholds) at specific stressor levels where the mean response values above and below the stressor level are significantly different (e.g., higher end of the shaded box in Fig. 2A). The analysis produces a regression tree in which clusters of observations grouped by stressor levels have different mean values. Lowess regression is a locally weighted regression based on sets of similar stressor values (Fig. 2B, C).

We did regression tree analysis with 1000 bootstrap regression tree simulations with the rpart and bootstrap routines in the statistical program R (R Core Development Team 2006) (King and Richardson 2003, Lougheed et al. 2007). We determined the probability that a threshold would be observed by chance and the range of TP concentrations within which the probability of observing a threshold was 0.10 to 0.90. We did lowess regression with S-PLUS (version 6; Insightful Corporation, Seattle, Washington; tension = 0.5), plotted the data with SYSTAT, and determined lowess thresholds as rules illustrated in Fig. 2 and described in the next paragraph. The likelihood that observed thresholds could occur by chance was evaluated by comparing the variance explained (r^2) by regression tree, lowess, and linear models in outputs from SYSTAT and S-PLUS, which in all cases was equivalent to the regression sum of squares divided by the total sum of squares. We assumed relationships were nonlinear if we observed a narrow range for TP concentrations within which the probability of observing a regression tree threshold was 0.10 to 0.90 or we observed greater r^2 for lowess and regression tree models than for linear models.

Lowess regression and regression tree analysis delineate thresholds or change-points in nonlinear responses of attributes to TP (Fig. 2A–C). We define *thresholds* as either a nonlinear response or the nutrient concentration at which the response occurs. Thresholds provide justification for narrowing the range of possible stressor levels that should be considered for criteria. Stressor criteria should be established at stressor levels below thresholds in nonlinear responses that demonstrate assimilative capacity (the ability of the system to resist change in the valued ecological attribute up to some level of the stressor) (Fig. 2A, B). Setting criteria below thresholds in responses demonstrating assimilative capacity provides a margin of safety to protect valued attributes, which is a requirement for setting total maximum daily loads (US Clean Water Act Section 303(d)(1)(c); USEPA 1999). Negative exponential responses with saturation levels (Fig. 2C) call for protecting attributes, processes, or condition before they are completely lost. Setting criteria at stressor levels below the saturation level (upper threshold) reduces risk of losing those attributes. The range of possible stressor criteria is widest with linear responses (Fig. 2D) when little rationale exists for specific criteria based on degree of loss of valued attributes with incremental increases in stressors. When responses are linear, criteria need to be established using other decision processes that are not evaluated here.

Results

Diatom taxonomic composition in reference streams indicated that factors associated with conductivity and alkalinity had strong effects on expected taxonomic composition that might interfere with our ability to use a 1-class reference-condition approach for setting TP criteria (i.e., species composition varied along a confounding natural gradient). One hundred nine taxa had an average relative abundance >1.0% in samples from a set of 123 reference sites. Two groups of sites were identified by TWINSPAN. One group was characterized by frequent occurrence of diatoms in the genus *Eunotia*, which are commonly found in low-alkalinity waters (van Dam et al. 1994). The other group was dominated by taxa common in alkaline waters. Discriminant functions (based on both forward and backward selection) identified dissolved organic C (DOC), conductivity, alkalinity, Si, and SO_4 as important factors affecting species composition in reference streams. Therefore, we used only well-buffered streams in the following analyses to control for natural variability among streams caused by pH. We defined well-buffered streams as those streams that showed little evidence of effects of mining and that met criteria for 2 discriminant functions based on forward and backward selection: ANC >50 and SO_4 <400 μeq/L (Herlihy et al. 1991). More streams in the data set were classified as well buffered than as poorly buffered. The TP benchmarks that are developed with these analyses would only apply to well-buffered streams.

Most molar N:P ratios in the well-buffered streams were greater than the 16:1 Redfield ratio, and these N:P values suggested that algae in most MAH streams would be P-limited if nutrient concentrations were low (Redfield 1958). Ratios ranged from 5 to 456 (median = 73, mean = 102, SD = 81). The 25th and 75th percentiles of the N:P distribution were 42 and 140.

TP

TP concentrations ranged from 1 to 103 μg/L (Fig. 3A, B). The median and 25th percentile of TP concentrations at all sites were 14 and 10 μg/L, respectively (Fig. 3A). TP concentrations in streams

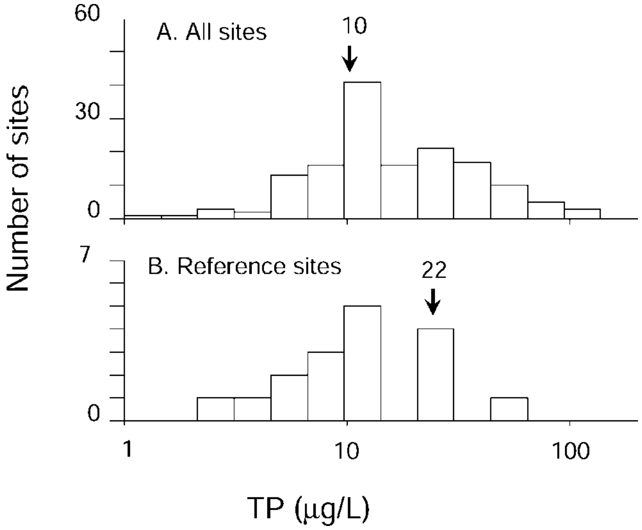

FIG. 3. Frequency distributions of total P (TP) concentrations at all sites (A) and sites in reference condition (B) in streams in the Mid-Atlantic Highlands region. Reference condition was defined as 0% watershed alteration by human land use. Arrows mark the 25th percentile of all sites in panel A and 75th percentile of reference sites in panel B. Numbers above the arrows are the TP associated with the percentile.

were positively related to % WA ($\log_{10}[\mathrm{TP}] = 2.265 + 0.017[\%\ \mathrm{WA}]$; $r^2 = 0.134$, $p < 0.001$). TP concentrations increased from a predicted mean of 9.6 μg TP/L at 0% WA to as high as 52.7 at 100% WA. The standard error (SE) of the y-intercept was 0.089; thus, the 75% and 90% CIs for predicted mean TP at 0% WA were 10.7 and 11.5 μg/L for streams with 0% WA.

The median and 75th percentiles of the TP distributions at reference sites with 0% WA ($n = 16$) were 10 and 22 μg/L, respectively (Fig. 3B). The number of sites with 0% WA was low. Therefore, we increased the maximum % WA in the definition of reference condition to 3% WA to increase sample sizes. The 75th percentile of the TP distribution was 12 μg/L at 3% WA (Fig. 4A), and did not increase above 12 μg/L until the maximum % WA for reference sites >10% ($n = 44$). The 75th percentile of the TP distribution increased much more rapidly than median TP concentration as we increased the maximum % WA in a sample set above 10% (Fig. 4A). The median and 75th percentiles of the TN (Fig. 4B) and Cl (Fig. 4C) distributions in streams also increased with % WA ($p < 0.001$). We did not observe lags in responses of TN and Cl concentrations to increasing % WA in low-disturbance watersheds (Fig. 4B, C).

Relationships of algal attributes to TP

Chl a *and AFDM.*—Ten of 11 algal attributes (all except chl *a*) were significantly linearly related to TP, but the explanatory value (r^2) of many of these models was low (Table 1). Benthic algal chl *a* and AFDM were highly variable among streams and ranged from <5 mg chl *a*/m^2 and 3 g AFDM/m^2 to >100 mg chl *a*/m^2 and 50 g AFDM/m^2 (Fig. 5A, B). In linear models, chl *a* and AFDM were only weakly related to TP concentrations. Among nonlinear models, regression tree models for ch *a* and AFDM were statistically signifi-

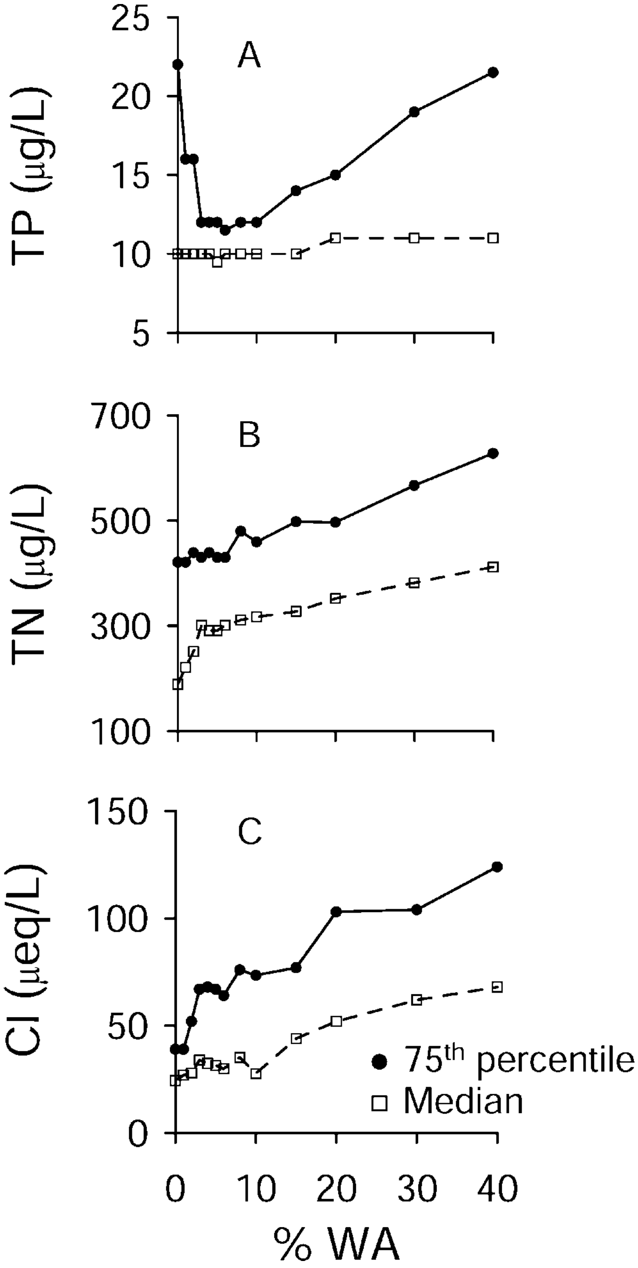

FIG. 4. Medians and 75th percentiles of frequency distributions of total P (TP) (A), total N (TN) (B), and Cl (C) concentrations when reference sites were defined with increasing levels of % watershed alteration (% WA) by human land use.

TABLE 1. Regression statistics indicating the relationships between algal assemblage attributes and total P (TP) concentrations in streams of the Mid-Atlantic Highlands ecoregion. a is the y-intercept and b is the regression coefficient in a regression of the form $y = a + b(x)$, N is the number of streams in the analysis, p is the attained significance (2-tailed t-test) of b, SE is the standard error of a or b. r^2 is the variance explained by linear, regression tree, and lowess regression. r^2 values in bold indicate $p < 0.05$ for the regression.

	Linear regression parameters								
Algal attribute	a	SE(a)	b	SE(b)	p	r^2	Regression tree r^2	Lowess r^2	N
ln(chlorophyll a)	2.898	0.327	0.196	0.113	0.087	0.023	**0.069**	0.054	127
ln(algal ash-free dry mass)	1.487	0.277	0.358	0.124	0.027	**0.040**	**0.068**	0.084	124
ln(acid phosphatase activity)	8.812	0.337	−0.289	0.116	0.014	**0.054**	**0.168**	**0.099**	112
ln(alkaline phosphatase activity)	7.632	0.384	−0.368	0.131	0.006	**0.064**	**0.134**	0.079	116
Number of observed taxa	15.727	3.079	4.255	1.074	<0.001	**0.097**	**0.143**	**0.147**	148
Evenness	0.488	0.036	0.05	0.012	<0.001	**0.100**	**0.136**	0.112	148
Proportion of native taxa	0.674	0.088	0.105	0.031	<0.001	**0.075**	**0.151**	**0.129**	148
Proportion low-P native taxa	1.119	0.082	−0.075	0.028	0.009	**0.045**	**0.060**	0.063	148
Taxonomic similarity to reference	0.488	0.029	−0.056	0.014	<0.001	**0.167**	**0.237**	**0.101**	148
% low-P individuals	56.593	4.942	−7.823	1.724	<0.001	**0.124**	**0.250**	0.138	148
ln(% high-P individuals)	−1.518	0.351	1.044	0.122	<0.001	**0.333**	**0.315**	**0.407**	140

cant, but lowess models for chl a and AFDM were not statistically significant (Table 1). In linear models, acid- and alkaline-phosphatase activities were significantly negatively related to TP concentrations in streams (Table 1, Fig. 5C, D). The regression-tree and lowess models of the relationship of acid phosphatase with TP were statistically significant, and the regression-tree model of the relationship between alkaline phosphatase and TP was statistically significant (Table 1).

Taxonomic composition and diversity.—Three hundred forty-five diatom taxa were observed in 193 well-buffered streams. On average, *Achnanthidium minutis-*

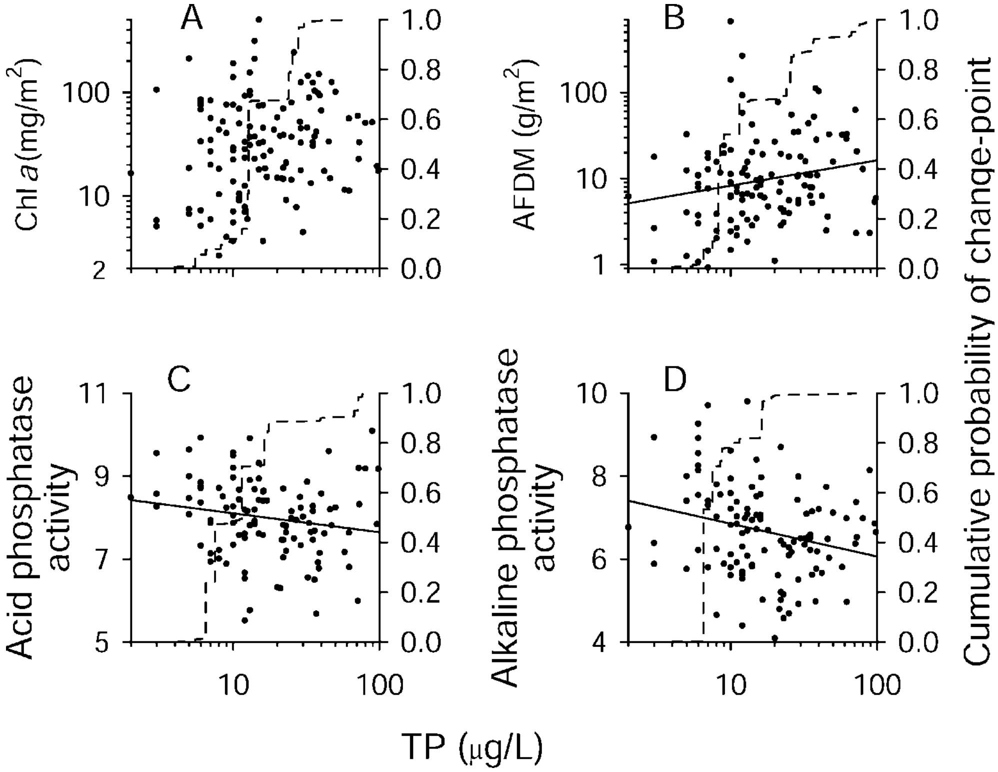

FIG. 5. Algal biomass measured as chlorophyll a (chl a) (A) and ash-free dry mass (AFDM) (B) and acid (C) and alkaline (D) phosphatase activity (nmoles activity g^{-1} AFDM h^{-1}) in Mid-Atlantic Highlands streams as a function of total P (TP) concentration. Solid lines indicate statistically significant predicted linear regressions ($p < 0.05$). Dashed lines indicate the cumulative probability of a change-point determined by bootstrapped regression tree analyses (see *Effects-based TP benchmarks* for details).

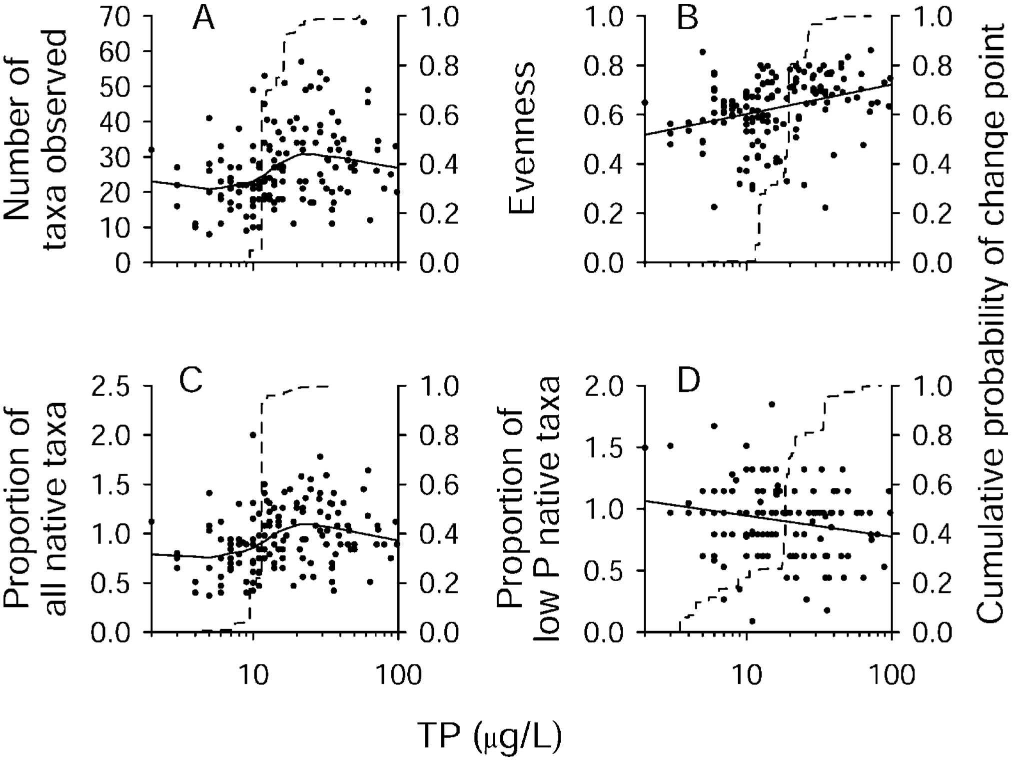

FIG. 6. Algal diversity measured as number of observed taxa (A), evenness (B), proportion of all native taxa (C), and proportion of low-P native taxa (D) (see *Effects-based TP benchmarks* for definitions and details of calculation of metrics) in Mid-Atlantic Highlands streams as a function of total P (TP) concentration. Solid lines indicate statistically significant predicted linear regressions ($p < 0.05$). Linear regressions are replaced by a smoothing line generated by lowess regression if the lowess regression explained 20% more variation in the diversity measure–TP relationship than did linear regression and r^2 of the lowess regression $\geq$0.10.

simum (Kützing) Czarnecki and *Achnanthes biasolettiana* (Kützing) Grunow constituted 30 and 20%, respectively, of assemblages in streams with TP <10 μg/L ($n = 36$ streams). One hundred forty taxa were observed in these low-TP streams, and only 7 taxa had an average regional relative abundance >2%. On average, *A. minutissimum* and *A. biasolettiana* constituted 14 and 11%, respectively, of assemblages in streams with TP >30 μg/L ($n = 35$ streams). However, 245 taxa were observed in these higher-TP streams, and 12 taxa had average regional abundances >2%.

Indicators of diversity (number of taxa observed and evenness) increased with TP, unless the taxa considered were low-P native taxa (Table 1, Fig. 6A–D). The number of diatom taxa observed (Fig. 6A) and the proportion of native taxa (Fig. 6C) increased significantly with TP in linear, regression tree, and lowess models, and evenness (Fig. 6B) increased significantly with TP in linear and regression tree models (Table 1). However, the proportion of low-P native taxa decreased significantly with increasing TP in linear and regression tree models (Table 1, Fig. 6D). Linear models explained only a small part of the variation in diversity metrics as a function of TP ($r^2 = 0.045$–1.00). Nonlinear models ($r^2 = 0.060$–0.151) explained substantially more variation than linear models for the number of taxa observed and proportion of native taxa, but not for evenness and proportion of low-P native taxa (Table 1). Lowess curves for number of taxa observed (Fig. 6A) and proportion of native taxa (Fig. 6C) indicated that the greatest change in numbers of taxa occurred between 8 and 20 μg TP/L.

Models relating changes in species composition to TP concentrations were more precise than models of biomass, diversity, and phosphatase activity as a function of TP (Table 1). Similarity to reference condition decreased from >0.45 to ~0.25 as TP increased from 0 to 100 μg/L (Fig. 7A). Percent low-P individuals in assemblages decreased from 40 to 50% at low TP concentrations to <30% in high TP concentrations (Fig. 7B). Percent high-P individuals increased from close to 0% at low TP concentrations to >20% at high TP concentrations (Fig. 7C). Linear and nonlinear models differed in their ability to explain these variables as a function of TP. The regression tree model explained more variability in similarity to reference than did the lowess model, and the linear model was intermediate between the 2 nonlinear models (Table 1). The regression tree model explained substantially more variability in % low-P individuals than did the linear model, but the lowess model was not statistically significant (Table 1). However, the lowess model explained substantially more variability in % high-P individuals than did either the linear or regression tree models (Table 1). Lowess regression indicated that similarity to reference conditions and % high-P individuals changed most rapidly at TP concentrations between 10 and 35 µg/L (Fig. 7A, C).

TP thresholds.—Most TP benchmarks derived by regression tree and lowess regression were between 10 and 20 µg/L. Lowess regression detected statistically significant nonlinear patterns for 5 of the 11 algal attributes (Table 1). Four of those 5 (number of observed taxa, proportion of native taxa, similarity to reference, % high-P individuals) showed threshold patterns with no response in biological condition when TP was <10 µg/L (Figs 6A, B, 7A, C). Percent high-P individuals increased from 1 to 3% at TP values <5 µg/L (Fig. 7C), but only 2 streams had very low TP concentrations and very low percentages of high-P individuals.

The most precise TP change-points derived from bootstrapped regression tree analyses were for the 4 algal attributes that had threshold patterns with lowess regression. The bootstrapped thresholds for number of taxa observed, proportion of native taxa, and % high-P individuals were 11.5 µg TP/L (Table 2). However, the bootstrapped threshold for similarity to reference was 26.5 µg TP/L (Table 2). The similarity-to-reference threshold change-point indicates one problem with regression tree analysis: >1 change-point might be identified in some threshold patterns, and the change-point identified might be at the higher end of the range of changing conditions. The change-points for these 4 variables were the most precise because these variables had the highest r^2, greatest number of observations of the same bootstrapped change-point, and the lowest deviation in TP between the 10th and 90th percentiles of the frequency distribution of bootstrapped change-points (Tables 1, 2).

We found some evidence for a threshold in chl *a*; the probability of chl *a* being >10 mg/m^2 was greater when TP was >12.7 µg/L (Fig. 5A). r^2 for the chl *a* regression tree analysis was only 0.069 (Table 1), but the 369 bootstrapped observations of the change-point at 12.7 increased to 510 if we added bootstrapped change-points recorded for just one stream with TP <12.7. Thus, the probability of observing chl *a* <10 mg/m^2 was low when TP was >12.7 (Fig. 5A). The relatively wide range in TP between 10th and 90th percentiles of the bootstrapped change-points was related to a 2nd change-point at ~21 µg TP/L.

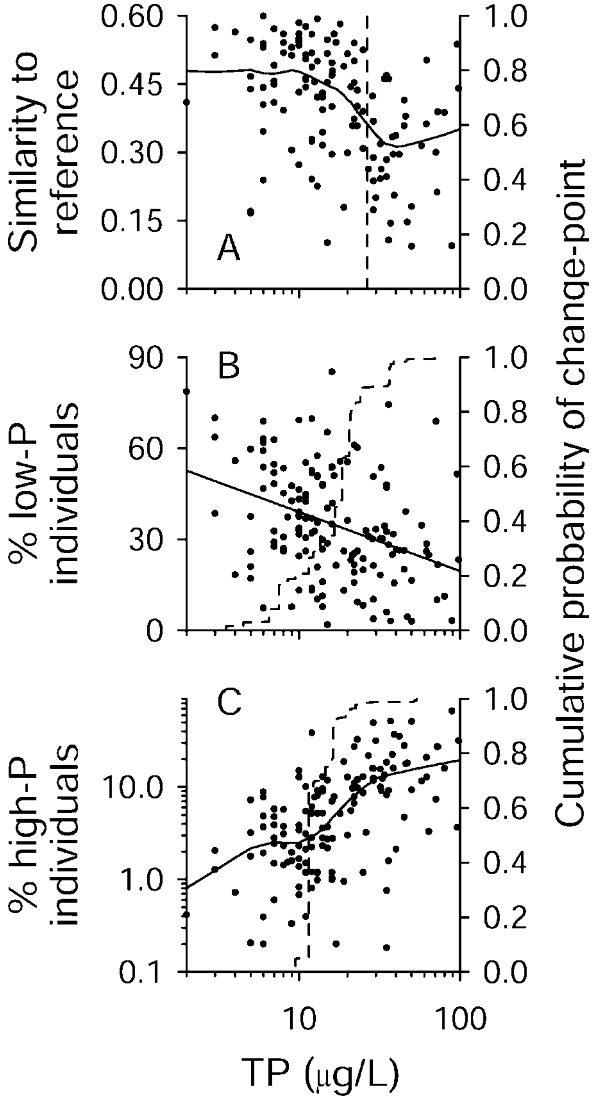

FIG. 7. Taxonomic composition of algal assemblages measured as the similarity of the assemblage at a site to reference conditions (A), % low-P individuals (B), and % high-P individuals (C) (see *Effects-based TP benchmarks* for definitions and details of calculation of metrics) in Mid-Atlantic Highlands streams as a function of total P (TP) concentration. Solid lines indicate statistically significant predicted linear regressions ($p < 0.05$). Linear regressions are replaced by a smoothing line generated by lowess regression if the lowess regression explained 20% more variation in the species composition measure–TP relationship than did linear regression and r^2 of the lowess regression ≥0.10.

Discussion

Expected condition and responses to P enrichment

The natural TP concentrations in MAH streams averaged ~10 µg/L, a value similar to those observed in low-nutrient lakes and wetlands (e.g., McCormick et al. 1996, Dixit et al. 1999) and in streams in watersheds

TABLE 2. Total P (TP) concentration (μg/L) benchmarks for nutrient criteria derived from biological responses (attributes) to TP using bootstrapped regression tree change-points (BCP) and lowess regression. 10th and 90th are the TP concentrations for the 10th and 90th percentiles of the cumulative frequency distribution of bootstrapped change-points. DLTP is the deviation in ln(TP) between the 10th and 90th percentiles, BCP is the bootstrapped change-point observed most often, and Boot no. is the number of bootstrapped observations (of 1000) that identified BCP. High Boot no. and low DLTP indicate more precise BCPs. NA indicates $p > 0.05$ or $r^2 > 0.10$ for lowess regression.

	Regression tree					
Algal attribute	10th	90th	BCP	Boot no.	DLTP	Lowess
ln(chlorophyll *a*)	8.5	27.8	12.7	369	0.51	NA
ln(algal ash-free dry mass)	7.5	36.5	8.2	261	0.69	NA
ln(acid phosphatase activity)	6.5	41	6.5	225	0.80	NA
ln(alkaline phosphatase activity)	6.5	16.3	6.5	531	0.40	NA
Number of observed taxa	11.5	16.5	11.5	619	0.16	10
Evenness	12.2	26.5	19.5	305	0.34	NA
Proportion of native taxa	9.5	11.5	11.5	709	0.08	10
Proportion low-P native taxa	4.5	34.5	18.5	217	0.88	NA
Species similarity to reference	26.5	26.5	26.5	1000	0.00	9
% low-P individuals	7.5	36.5	18.5	154	0.69	NA
ln(% high-P individuals)	11.5	16.5	11.5	619	0.16	10

with low levels of human land use in other regions of the US (Stevenson et al. 2006). Regression analysis predicted that natural TP was 9.6 μg/L and that median TP concentrations were 10 μg/L when human disturbance in watersheds was 0% WA. However, this benchmark might vary among regions despite the fact that natural nutrient concentrations are ~10 μg/L in aquatic ecosystems of many regions. Dodds and Oakes (2004) predicted that variation in natural TP concentrations would range from 20 to >100 μg/L among ecoregions. Most variability in TP concentrations among streams in the MAH was associated with human disturbance in watersheds, as noted elsewhere (Allan et al. 1997, Dodds and Oakes 2004), but distinguishing between natural and anthropogenic sources of P is essential for assessing and managing human effects on ecosystems.

Differences in benthic algal biomass, enzyme activity, diatom diversity, and diatom species composition were related to TP concentrations in MAH streams. Some of these observations are common in studies of algal assemblages, but all of these attributes are seldom measured in the same study and in hundreds of streams. In general, the precision of relationships increased with taxonomic resolution; indices based on species composition were much more precisely related to TP ($r^2 \approx 0.12$–0.41) than were assemblage-level measures of biomass ($r^2 \approx 0.02$–0.08).

The relationship between algal biomass and nutrient concentrations in stream water is notoriously complex (Stevenson 1997, Dodds et al. 1998, Biggs 2000). Time elapsed since scouring by storms, shading, grazers, patchiness caused by macroalgal colonization, and current velocity can affect algal biomass and its estimation in streams (Stevenson et al. 1996, Kohler and Wiley 1997, Biggs 2000, Riseng et al. 2004). Moreover, benthic algae remove nutrients from the water column (Mulholland 1996), and this removal can mask the relationship between benthic algae and limiting nutrients. In our study, use of nonlinear regression techniques increased r^2 for chl *a* and AFDM relationships with TP, but TP explained less variation in algal biomass in our study than in others (Dodds et al. 1997, 2002, Stevenson et al. 2006). Diverse hydrogeomorphology of MAH streams, seasonal and spatial variability in biomass, a single measurement of nutrient concentrations, and negative effects of biomass on soluble P concentration probably constrained our ability to discern precise relationships between biomass and TP.

The activities of both acid and alkaline phosphatase decreased with increasing TP in our study. We are unaware of any other large-scale studies in streams that showed decreases in phosphatase activities as nutrients increased. However, phosphatase activity is seldom measured in large-scale surveys. Phosphatase activity decreased with P enrichment at TP concentrations >10 to 15 μg/L in the Everglades, a large freshwater marsh (Richardson and Qian 1999). The relationships between phosphatase activity and TP are highly variable and are influenced by many factors. For example, algal biomass could confound our understanding of phosphatase–TP relationships if high biomass generated nutrient limitation within the algal mat. Nutrient limitation might occur if the rates of nutrient uptake within a mat were higher than the rate of nutrient transport into the mat (Stevenson and Glover 1993). Nevertheless, in our study, the decrease in phosphatase activity as TP increases suggests that P limitation in MAH streams is relieved as TP increases.

The number of observed taxa and evenness increased as TP increased. This result probably reflects a release of algae from severe P limitation and increased habitat suitability for high-P taxa. The number of observed diatom taxa in counts is a poor estimate of the true number of species in a habitat when only 500 valves (halves of individual diatom cell walls) and a small fraction of species in samples are observed (Patrick et al. 1954, Stevenson and Lowe 1986). The number of observed taxa in such counts is highly correlated with evenness (Archibald 1972, Stevenson and Lowe 1986). In our study, the number of observed taxa also was correlated with evenness because high-P taxa were added to low-P assemblages and more taxa had relative abundances >2% as TP increased.

Growth rates of low-P taxa increase more slowly than growth rates of high-P taxa when P concentrations increase, and this phenomenon might explain the increase in evenness with increasing P concentrations in streams at the lower end of the TP range. Shifts from low- to high-P taxa with increasing TP would result if growth rates of low-P taxa became saturated at lower TP concentrations than did growth rates of high-P taxa, and if low-P taxa had slower maximum growth rates than high-P taxa. Experimental results confirm that *A. minutissimum* and *A. biasolettiana*, the 2 dominant taxa in low-P MAH streams, have relatively high growth rates in low-P conditions, but that other taxa have higher growth rates than do *A. minutissimum* and *A. biasolettiana* in high-P conditions (Manoylov and Stevenson 2006). Thus, as TP concentrations increase, species growth rates become more similar, relative abundances become more even, and more taxa are observed in a 500-valve count. In MAH streams, % high-P individuals increased as TP increased. Thus, invasion of assemblages by diatom species that require high resource levels seems to be prevented in near-natural MAH streams by low P availability. If a low-P habitat cannot sustain high-P taxa, then release from resource constraints (i.e., addition of P) might increase taxon numbers. However, estimates of species richness based on 500-valve counts are not sufficient to test this hypothesis.

We found some evidence that low-P native species might be negatively affected by increased TP. The number of low-P native taxa and % low-P individuals decreased with increasing TP. The r^2 value of the relationship between the number of low-P native taxa and TP was low, but the 25% reduction in % low-P individuals as TP increased might be ecologically significant if it indicates a proportional loss of low-P native taxa. Loss of native taxa as resources increase could be caused by one of many possible density-dependent biotic interactions, such as competition for other resources, allellopathy, habitat alteration, or viral or bacterial attack. The decrease in proportion of expected low-P native species in samples could indicate a negative effect of nutrient enrichment on biodiversity of native microbial taxa.

Candidate nutrient criteria

Multiple lines of evidence indicate that a nutrient criterion of 10 to 12 μg/L can be justified for MAH wadeable streams based on reference conditions and benthic algal responses to nutrients. Natural concentrations of TP were predicted to be 9.6 μg/L with a 75% CI of 10.7 μg/L at 0% WA. The median and 75th percentile of TP concentrations at reference sites (WA < 10%) were 10 and 12 μg/L, respectively.

Increases in TP from 10 to 30 μg/L were associated with responses of benthic diatom assemblages that indicated release from nutrient limitation. Nonlinear regressions showed that some biological variables differed little among streams when TP was <10 μg/L, but differed greatly as TP increased from 10 to 30 μg/L. Observational and experimental studies have shown that benthic algae in streams respond to changes in PO_4-P or TP between 1 and 30 μg/L (e.g., Bothwell 1989, Dodds et al. 1997, Chételat et al. 1999, Rier and Stevenson 2006, Stevenson et al. 2006). Algae in wetlands and lakes respond to changes in TP at concentrations as low as 10 μg/L (McCormick et al. 1996, Dixit et al. 1999). Thus, many attributes of benthic algal assemblages are sensitive to increases in nutrient concentrations above natural background levels.

Protection of all MAH streams with a TP criterion of 10 to 12 μg/L probably is not practical. We did not observe strong evidence for using other TP criteria, but others have. Extensive, long-lasting nuisance growths of the filamentous green alga *Cladophora* that are unaesthetic and a problem for fishing and recreational use can be prevented by maintaining an average of 30 μg TP/L (Dodds et al. 1997, Stevenson et al. 2006). Effect of nutrients on fish assemblages occur when TN > 0.61 mg/L and TP > 0.06 mg/L (Miltner and Rankin 1998, Wang et al. 2007). Any of these benchmarks (i.e., concentrations at which TP causes important ecological responses) could be used as a TP criterion that would provide some desired level of protection for streams and support for management goals corresponding to *tiered aquatic life uses* (Davies and Jackson 2006). Within a region, different nutrient criteria could be identified and applied selectively to streams that have different management goals. For example, low stressor criteria, such as 10 μg TP/L, would prevent degradation of high-quality sites and set targets for restoration that would correspond to the goals of biological integrity in the US Clean Water Act. A 10-μg TP/L criterion would constrain productivity to natural levels, maintain P-resource constraints on species membership, and prevent proliferation of invasive algae that require high nutrient concentrations. On the other hand, higher stressor criteria, such as between 30 and 60 μg TP/L, could be interpreted as protective of "fish, shellfish, and wildlife" that would correspond to the lower, interim goal of the US Clean Water Act.

The different approaches for defining nutrient criteria used in our study provided complementary information that should aid criterion development. Predictive models confirmed that human activities increased TP in streams and estimated the TP concentration that should be expected when human activities with watersheds are low. Frequency distributions of TP at sites with low WA characterized minimally disturbed conditions. Frequency distributions of TP at all sites described the existing range of TP concentrations within the region and provided an indication of proportion of streams that would require restoration for each TP benchmark. Relationships between variables indicating biological condition and TP described the valued attributes that could be supported at successively higher TP concentrations. Change-points in stressor–response relationships provided important benchmarks along the TP gradient that justified special attention (as suggested by Muradian 2001 and as observed in our discussions with stakeholders).

The degree to which our results can be transferred from MAH streams to other streams and other types of aquatic ecosystems will depend on the analytical approach and the ecosystem. Regional factors, such as geochemistry, climate, and hydrology, and local factors, such as stream size and gradient, will affect the expected natural TP concentrations predicted by models or the frequency distributions of reference conditions (e.g., Dodds and Oakes 2004). Regional and local factors also might affect responses of algae and invertebrates to increases in TP (Biggs 1995, Riseng et al. 2004). Characterizations of minimally disturbed conditions with frequency distributions also will vary greatly with the extent of human alteration of regions. Two of our approaches (evaluation of frequency distributions of reference sites and thresholds in stressor–response relationships) identified TP benchmarks at 10 to 12 μg TP/L. This co-occurrence probably is related to the sensitivity of algae to deviations from natural conditions and to the substantial proportion of MAH streams with low levels of human disturbance. Had we studied a region with more extensive alteration of watersheds than MAH, we would have been able to estimate natural conditions with predictive models as long as some watersheds had % WA ≈ 0; however, we might not have had sufficient sample size to characterize central tendency and variability in natural conditions precisely with frequency distributions. Nevertheless, if climate, geology, and stream hydrogeomorphology are sufficiently similar, we expect transferability of results from regression-based stressor–response approaches and threshold benchmarks. For example, responses of *Cladophora* in riffles to P are very similar in Montana and the Midwest (Dodds et al. 1997, Stevenson et al. 2006).

In conclusion, multiple analytical approaches can and should be used when developing nutrient criteria to provide the diversity of information that justify criteria to stakeholders and increase the probability of successful management actions. These approaches have successfully characterized best available conditions and effects of nutrients on algae. However, variations in nutrient concentrations and responses probably are related to natural and anthropogenic factors in complex ways that call for future work to characterize how predicted best available condition and ecological responses to nutrients vary among regions. In addition, experiments in which nutrients are manipulated are needed to confirm cause–effect relationships, especially with nuisance macroalgae in streams. The complementary use of predictive models, frequency distributions, and stressor–response relationships provides a sound analytical basis for criterion development by characterizing best available conditions and quantifying effects on valued ecological attributes.

Acknowledgements

The Office of Research and Development of the US EPA supported the work done by RJS on this paper. Robert Hughes and members of the US EPA Tiered Aquatic Life Use Workgroup provided valuable conversations in the formulation of the concepts evaluated in this paper. Patti Grace-Jarrett, Kalina Manoylov, Mark VanderBorgh, and Susanna Decelles identified and counted algae.

Literature Cited

ALLAN, J. D., D. L. ERICKSON, AND J. FAY. 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater Biology 37:149–161.

APHA (American Public Health Association). 1992. Standard methods for the evaluation of water and wastewater. 18th edition. American Public Health Association, American Water Works Association, and Water Environment Federation, Washington, DC.

Archibald, R. E. M. 1972. Diversity in some South African diatom assemblages and its relation to water quality. Water Research 6:1229–1238.

Biggs, B. J. F. 1995. The contribution of disturbance, catchment geology and land use to the habitat template of periphyton in stream ecosystems. Freshwater Biology 22:209–231.

Biggs, B. J. F. 2000. Eutrophication of streams and rivers: dissolved nutrient–chlorophyll relationships for benthic algae. Journal of the North American Benthological Society 19:17–31.

Biggs, B. J. F., and R. A. Smith. 2002. Taxonomic richness of stream benthic algae: effects of flood disturbance and nutrients. Limnology and Oceanography 47:1175–1186.

Bothwell, M. L. 1989. Phosphorus-limited growth dynamics of lotic periphytic diatom communities: areal biomass and cellular growth rate responses. Canadian Journal of Fisheries and Aquatic Sciences 46:1293–1301.

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8:559–568.

Chételat, J., F. R. Pick, A. Morin, and P. B. Hamilton. 1999. Periphyton biomass and community composition in rivers of different nutrient status. Canadian Journal of Fisheries and Aquatic Sciences 56:560–569.

Davies, S. P., and S. K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16:1251–1266.

Dixit, S. S., J. P. Smol, D. F. Charles, R. M. Hughes, S. G. Paulsen, and G. B. Collins. 1999. Assessing water quality changes in the lakes of the northeastern United States using sediment diatoms. Canadian Journal of Fisheries and Aquatic Sciences 56:131–152.

Dodds, W. K., J. R. Jones, and E. B. Welch. 1998. Suggested classifications of stream trophic state: distributions of temperature stream types by chlorophyll, total nitrogen, and phosphorus. Water Research 32:1455–1462.

Dodds, W. K., and R. M. Oakes. 2004. A technique for establishing reference nutrient concentrations across watersheds affected by humans. Limnology and Oceanography: Methods 2:331–341.

Dodds, W. K., V. H. Smith, and K. Lohman. 2002. Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. Canadian Journal of Fisheries and Aquatic Sciences 59:865–874.

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.

Francoeur, S. N. 2001. Meta-analysis of lotic nutrient amendment experiments: detecting and quantifying subtle responses. Journal of the North American Benthological Society 20:358–369.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications 10:1456–1477.

Herlihy, A. T., P. R. Kaufmann, and M. E. Mitch. 1991. Chemical characteristics of streams in the Eastern United States: II. Sources of acidity in acidic and low ANC streams. Water Resources Research 27:629–642.

Herlihy, A. T., J. L. Stoddard, and C. B. Johnson. 1998. The relationship between stream chemistry and watershed land cover data in the Mid-Atlantic region, U.S. Water, Air, and Soil Pollution 105:377–386.

Hill, B. H., A. T. Herlihy, P. R. Kaufmann, R. J. Stevenson, and F. H. McCormick. 2000. The use of periphyton assemblage data as an index of biotic integrity. Journal of the North American Benthological Society 19:50–67.

Hill, M. O., R. G. H. Bounce, and M. W. Shaw. 1975. Indicator species analysis, a divisive polythetic method of classification and its application to a survey of native pinewoods in Scotland. Journal of Ecology 63:597–613.

Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577–586.

Huston, M. A. 1994. Biological diversity: the coexistence of species on changing landscapes. Cambridge University Press, Cambridge, UK.

Huston, M. A. 2004. Management strategies for plant invasions: manipulating productivity, disturbance, and competition. Diversity and Distributions 10:167–178.

Juggins, S., and C. J. F. ter Braak. 1992. CALIBRATE: a program for species–environment calibration by [weighted averaging] partial least squares regression. Environmental Change Research Center, University College, London, UK. (Available from: http://www.campus.ncl.ac.uk/staff/Stephen.Juggins/software.htm)

Kaufmann, P. R., P. Levine, E. G. Robison, C. Seeliger, and D. V. Peck. 1999. Quantifying physical habitat in wadeable streams. EPA 620-R-99-003. Office of Research and Development, US Environmental Protection Agency, Corvallis, Oregon.

Kelly, M. G., and B. A. Whitton. 1995. The trophic diatom index: a new index for monitoring eutrophication in rivers. Journal of Applied Phycology 7:433–444.

King, R. S., and C. J. Richardson. 2003. Integrating bioassessment and ecological risk assessment: an approach to develop numeric water quality criteria. Environmental Management 31:795–809.

Kohler, S. L., and M. J. Wiley. 1997. Pathogen outbreaks reveal large-scale effects of competition in stream communities. Ecology 78:2164–2176.

Lazorchak, J. M., D. J. Klemm, and D. V. Peck. 1998. Environmental Monitoring and Assessment Program—surface waters: field operations and methods for measuring the ecological condition of wadeable streams. EPA-620-R-94-004F. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

Legendre, P., and L. Legendre. 1998. Numerical ecology. Elsevier Science, Amsterdam, The Netherlands.

Lougheed, V. L., C. A. Parker, and R. J. Stevenson. 2007. Using non-linear responses of multiple taxonomic groups to establish criteria indicative of wetland biological condition. Wetlands 27:96–109.

Manoylov, K. M., and R. J. Stevenson. 2006. Density-dependent algal growth along N and P nutrient gradients in artificial streams. Pages 333–352 *in* N. Ognjanova-Rumenova and K. Manoylov (editors). Advances in phycological studies. Pensoft Publishers, Moscow, Russia.

McCormick, P. V., P. S. Rawlick, K. Lurding, E. P. Smith, and F. H. Sklar. 1996. Periphyton–water quality relationships along a nutrient gradient in the northern Florida Everglades. Journal of the North American Benthological Society 15:433–449.

Miltner, R. J., and E. T. Rankin. 1998. Primary nutrients and the biotic integrity of rivers and streams. Freshwater Biology 40:145–150.

Mulholland, P. J. 1996. Role in nutrient cycling. Pages 609–639 *in* R. J. Stevenson, M. L. Bothwell, and R. L. Lowe (editors). Algal ecology, freshwater benthic ecosystems. Academic Press, San Diego, California.

Muradian, R. 2001. Ecological thresholds: a survey. Ecological Economics 38:7–24.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118–125.

Pan, Y., R. J. Stevenson, B. Hill, P. R. Kaufmann, and A. T. Herlihy. 1999. Spatial patterns and ecological determinants of benthic algal assemblages in Mid-Atlantic Highland streams. Journal of Phycology 35:460–468.

Pan, Y., R. J. Stevenson, B. H. Hill, A. T. Herlihy, and G. B. Collins. 1996. Using diatoms as indicators of ecological conditions in lotic systems: a regional assessment. Journal of the North American Benthological Society 15:481–495.

Pan, Y., R. J. Stevenson, P. Vaithiyanathan, J. Slate, and C. J. Richardson. 2000. Changes in algal assemblages along observed and experimental phosphorus gradients in a subtropical wetland, U.S.A. Freshwater Biology 44:339–353.

Patrick, R., M. H. Hohn, and J. H. Wallace. 1954. A new method for determining the pattern of the diatom flora. Notulae Nature 259:1–12.

R Core Development Team. 2006. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (Available from: http://www.r-project.org/)

Redfield, A. C. 1958. The biological control of chemical factors in the environment. American Scientist 46:205–221.

Richardson, C. J., and S. Qian. 1999. Long-term phosphorus assimilative capacity in freshwater wetlands: a new paradigm for maintaining ecosystem structure and function. Environmental Science and Technology 33: 1545–1551.

Rier, S. T., and R. J. Stevenson. 2006. Response of periphytic algae to gradients in nitrogen and phosphorus in streamside mesocosms. Hydrobiologia 561:131–147.

Riseng, C. M., M. J. Wiley, and R. J. Stevenson. 2004. Hydrologic disturbance and nutrient effects on benthic community structure in midwestern US streams: a covariance structure analysis. Journal of the North American Benthological Society 23:309–326.

Sayler, G. S., M. Puziss, and M. Silver. 1979. Alkaline phosphatase assay for freshwater sediments: application to perturbed sediment systems. Applied and Environmental Microbiology 12:1522–1539.

Setzer, R. W., and C. A. Kimmel. 2003. Use of NOAEL, benchmark dose, and other models for human risk assessment of hormonally active substances. Pure and Applied Chemistry 75:2151–2158.

Smith, V. H., G. D. Tilman, and J. C. Nekola. 1999. Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. Environmental Pollution 100:179–196.

Stevenson, R. J. 1997. Resource thresholds and stream ecosystem sustainability. Journal of the North American Benthological Society 16:410–424.

Stevenson, R. J., R. C. Bailey, M. C. Harass, C. P. Hawkins, J. Alba-Tercedor, C. Couch, S. Dyer, F. A. Fulk, J. M. Harrington, C. T. Hunsaker, and R. K. Johnson. 2004. Interpreting results of ecological assessments. Pages 85–111 *in* M. T. Barbour, S. B. Norton, H. R. Preston, and K. W. Thornton (editors). Ecological assessment of aquatic resources: linking science to decision-making. Society of Environmental Toxicology and Contamination, Pensacola, Florida.

Stevenson, R. J., M. Bothwell, and R. L. Lowe (editors). 1996. Algal ecology: freshwater benthic ecosystems. Academic Press, San Diego, California.

Stevenson, R. J., and R. Glover. 1993. Effects of algal density and current on ion transport through periphyton communities. Limnology and Oceanography 38:1276–1281.

Stevenson, R. J., and R. L. Lowe. 1986. Sampling and interpretation of algal patterns for water quality assessment. Pages 118–149 *in* B. G. Isom (editor). Rationale for sampling and interpretation of ecological data in the assessment of freshwater ecosystems. ASTM STP 894. American Society for Testing and Materials, Philadelphia, Pennsylvania.

Stevenson, R. J., Y. Pan, and P. Vaithiyanathan. 2002. Ecological assessment and indicator development in wetlands: the case of algae in the Everglades, USA. Verhandlungen der Internationalen Vereinigung für theoretische und andgewandte Limnologie 28:1248–1252.

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 2 regions with different disturbance regimes and with applications for developing nutrient criteria. Hydrobiologia 561:149–165.

Suter, G. W. 1993. Ecological risk assessment. CRC Press, Boca Raton, Florida.

USEPA (US Environmental Protection Agency). 1987. Handbook of methods for acid deposition studies, laboratory analysis for surface water chemistry. EPA-600-4-87-026. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

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). 1999. Protocol for developing nutrient TMDLs. EPA 841-B-99-007. Office of Water, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000. Nutrient criteria technical guidance manual. Rivers and streams. EPA-822-B-00-002. Office of Water, 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. Netherlands Journal of Aquatic Ecology 28:117–133.

VITOUSEK, P. M., J. D. ABER, R. W. HOWARTH, G. E. LIKENS, P. A. MATSON, D. W. SCHINDLER, W. H. SCHLESINGER, AND D. G. TILMAN. 1997. Human alteration of the global nitrogen cycle: sources and consequences. Ecological Applications 7:737–750.

WAITE, I. R., A. T. HERLIHY, D. P. LARSEN, AND D. J. KLEMM. 2000. Comparing the strengths of geographic and non-geographic classifications of stream benthic macroinvertebrates in the Mid-Atlantic Highlands, USA. Journal of the North American Benthological Society 19:429–441.

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

WRIGHT, J. F., M. T. FURSE, AND P. D. ARMITAGE. 1993. RIVPACS: a technique for evaluating the biological quality of rivers in the UK. European Water Pollution Control 3:15–25.

Received: 13 July 2007
Accepted: 10 June 2008