J. N. Am. Benthol. Soc., 2005, 24(2):415–427

Choice of substrate in algae-based water-quality assessment

MARINA POTAPOVA[1] AND DONALD F. CHARLES[2]

Patrick Center for Environmental Research, The Academy of Natural Sciences, 1900 Benjamin Franklin Parkway, Philadelphia, Pennsylvania 19103 USA

Abstract. Our study investigated whether algae-based water-quality assessments are affected by differences between algal assemblages on hard substrates (rocks, wood) and soft substrates (fine-grained sediments). We analyzed a US Geological Survey National Water-Quality Assessment (NAWQA) program data set that consisted of 1048 pairs of samples collected from hard and soft substrates at 551 river sampling locations throughout the US. Biovolume and diversity of algal assemblages, biovolume of major taxonomic groups, and abundance of motile diatoms differed significantly between samples collected from hard and soft substrates at the same sites. Ordinations of assemblages from hard and soft substrates were highly concordant and provided similar information on environmental gradients underlying species patterns. The strengths of relationships between composition of algal assemblages and water chemistry parameters (conductivity, pH, total P, and total N) did not differ consistently between substrate types. Performance of weighted averaging (WA) inference models did not differ between models based on assemblages from hard and soft substrates. Moreover, the predictive power of inference models developed from single-substrate data sets was not reduced when these models were applied to samples collected from other substrates. We concluded that the choice of substrate to sample should depend on the assessment indicators to be used. If indicators based on the autecologies of many algal taxa (e.g., inference models or autecological indices) are used, restricting samples to a single type of substrate is unnecessary. If algal diversity, total algal biovolume, or abundance of specific algal taxa is used, samples should be collected from a single type of substrate.

Key words: benthic algae, substrate, water quality, monitoring, rivers, inference models, indicator species, NAWQA, diatoms.

Species composition of benthic algal communities sampled at the same site, but from different substrates (e.g., rock surface, upper layer of sediment, or water plants), often differs substantially because some species are better adapted to one microhabitat than to others (Round 1981). The abundance of algae also varies with substrate type (Stevenson and Hashim 1989, Sabater et al. 1998). These natural between-substrate differences potentially confound responses of algal assemblages to stresses associated with human activities and may interfere with water-quality assessments based on knowledge of these responses.

The choice of substrate to sample is particularly important in large-scale water-quality assessment surveys that are carried out in diverse landscapes, where a single substrate type may not occur at all sampling sites. For example, rocks may be plentiful in high-gradient streams, but rare in low-gradient rivers. It is important to know whether values of metrics or indices calculated for samples collected from one substrate type can be compared with those for samples from another substrate type. If not, samples may have to be collected from multiple substrate types, and this strategy would require additional effort and expense. The purpose of our study was to provide information about the effect of substrate selection on water-quality assessments.

Previous algal studies do not provide clear guidance on whether sampling in water-quality assessment surveys should be restricted to a standard substrate. Studies estimating the effects of point-source pollution have shown that sampling the same substrate is necessary (Lowe and Pan 1996, Kelly et al. 1998). However, some surveys of large geographic areas have not found significant between-substrate differences in algal assemblage structure (Jüttner et al. 1996, Soininen and Eloranta 2004), possibly because the effects of other environmental factors were overriding. Other studies have shown that algal metrics/indices used in water-quality monitoring do not differ significantly when applied to samples collected from different types

[1] E-mail addresses: potapova@acnatsci.org
[2] charles@acnatsci.org

of substrate (Rott et al. 1998, Kitner and Poulíčková 2003).

US Geological Survey (USGS) National Water-Quality Assessment (NAWQA) program personnel have collected quantitative benthic algal samples from hard (rocks when present, otherwise wood) and soft (fine-grained sediments) substrates in rivers across the US (Gurtz 1993). NAWQA protocols recommend sampling both hard and soft substrates at all sampling sites (Porter et al. 1993, Moulton et al. 2002). We used the NAWQA national-scale data set of paired benthic algal samples to address the following questions: 1) Do algal biovolume, diversity, and species composition differ between hard and soft substrates at the national and ecoregional scales? 2) Do relationships between assemblage composition and water quality differ depending on the substrate sampled? 3) Is one substrate type better suited than others for developing water-quality assessment techniques based on the autecologies of many taxa? 4) Can water-quality assessment techniques developed for one substrate be applied to samples collected from another substrate?

Methods

Sample collection and laboratory analyses

Benthic algal samples were collected by USGS NAWQA program personnel between 1993 and 2001 from rivers throughout the US. Data from 551 sampling locations where pairs of samples were taken from 2 types of substrate were used for national-scale comparisons (Fig. 1). One substrate was fine-grained or soft sediment (*Depositional-Targeted Habitat* in NAWQA protocols). In some cases, this substrate was described as sand or silt, but often no distinction was made. The second substrate was hard (*Richest-Targeted Habitat* in NAWQA protocols). Hard substrates were rocks or submerged woody debris if rocks were not available. Additional (non-paired) NAWQA samples were used in the indicator species analysis and to test some inference models developed in our study. Samples were collected annually at each site for 1 to 3 y. Most algal samples were collected during low-flow conditions, usually in summer or early autumn. Samples from hard substrates were collected with an SG-92 sampling device, which is a modified plastic syringe with a sampling area of $\sim$3 cm^2 (Porter et al. 1993). At least 5 rocks or snags were sampled randomly at each sampling location, and 5 locations were sampled at each sampling reach (150–1000 m in length). These samples were pooled to form a single sample; the total sampled area was $\geq$75 cm^2. When algae were very sparse or the hard-substrate surface was too rough, algae were scraped and brushed from whole rocks (or sections of snags and sticks), and the area sampled was determined by the foil template method (Porter et al. 1993). Soft-sediment samples were collected from the top 5- to 7-mm layer of sediment by pushing a Petri dish lid (area = 17 cm^2) into the sediment surface and sliding a spatula underneath. Samples from 5 locations in each sampling reach were pooled into a single sample; the total sampled area was 85 cm^2.

Algal samples were analyzed at the Patrick Center for Environmental Research of The Academy of Natural Sciences, Philadelphia (ANSP), at J. R. Stevenson's laboratories at the University of Louisville and Michigan State University, and by independent contractors. Laboratory methods were described in Charles et al. (2002). Nondiatom algae were identified and counted in Palmer–Maloney counting chambers. Diatom cells containing plastids, and presumably alive when collected, were counted, but were not identified. Species-level identifications and counts of diatoms were made from material permanently mounted on microscope slides. The density (cells/cm^2) of each diatom species was calculated by multiplying the relative abundance of each species estimated from the permanent-mount diatom count by the total number of live diatoms (cells/cm^2) determined from the Palmer–Maloney count. This method overestimated live-diatom species richness because some of the dead cells undoubtedly originated from habitats other than the sampled substrate, and these allochthonous diatoms could not be distinguished from autochthonous diatoms in the permanent slide.

Environmental data, including site elevation, watershed area, channel gradient, and mean annual air temperature were provided by J. Falcone (USGS, Reston, Virginia). Water chemistry data used in our analysis were retrieved from the NAWQA Data Warehouse (http://water.usgs.gov/nawqa). We used measurements of total P (TP), total N (TN), and NO_3 concentrations, conductivity, alkalinity, and pH that had been

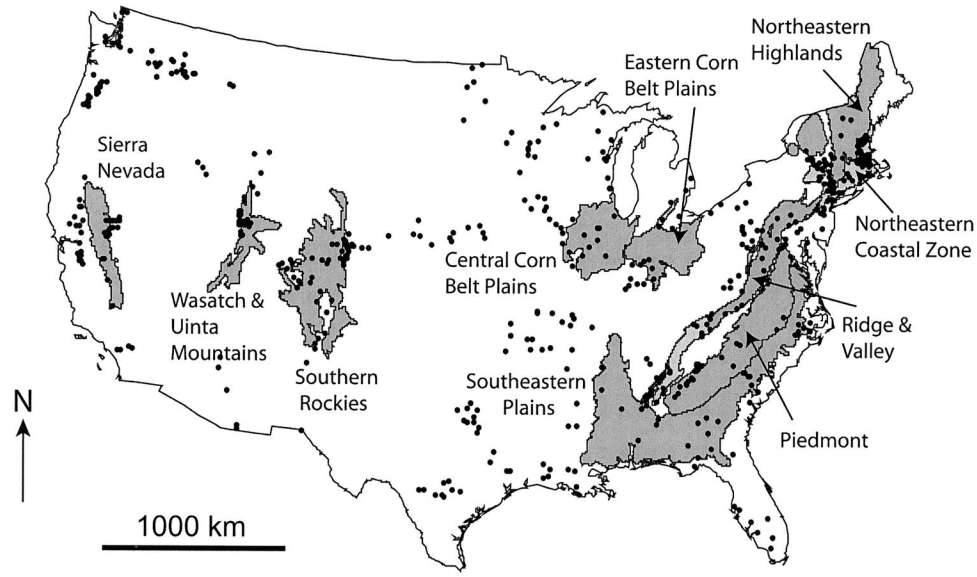

FIG. 1. Locations of 551 National Water-Quality Assessment sampling sites and 10 selected ecoregions (shaded). Samples from the Central Corn Belt Plains and Eastern Corn Belt Plains ecoregions were combined and treated as representing one ecoregion in the data analysis (see text).

made closest to the dates of algal sampling. NAWQA personnel collected algal and water chemistry samples within 1 mo of each other, and most (¾ of all samples) were collected within 2 wk of each other. We acknowledge that such time lags may have reduced our ability to infer relationships between algal assemblages and water chemistry. However, we think the potential effects of time lags were acceptable for the purposes of our study. Lag times varied among sites, but algal samples from both substrates at each site had the same lag time, and paired substrates were the basis for comparisons, not sites. Moreover, all sampling was confined to periods of stable low flows when within-site fluctuations of water chemistry were small compared to between-site variability.

Data analysis

Algal assemblages from different habitats were compared at the national scale, using a data set of 1048 pairs of samples collected from 551 sampling locations throughout the continental US, and the ecoregion scale, using 9 selected data sets. The data were analyzed at the ecoregion scale because water-quality monitoring tools and techniques often are developed at this scale (Barbour et al. 1999). Moreover, the ecoregion framework currently is used by most US federal and state environmental agencies for management of aquatic ecosystems and their components (Hughes and Larsen 1988). Level III ecoregions were used to create ecoregion-scale data sets. Level III is a hierarchical level of ecoregion classification in a scheme where Level I is the coarsest spatial scale and Level IV is the finest (Omernik 1987). Our intent was to include regions of the US with contrasting relief and climate. The western mountains were represented by the Sierra Nevada, Wasatch and Uinta Mountains, and Southern Rockies ecoregions; the eastern highlands were represented by the Ridge and Valley and Northeastern Highlands ecoregions; and the eastern coastal plains were represented by the Southeastern Plains, Piedmont, and Northeastern Coastal Zone ecoregions (Fig. 1). Each ecoregion was represented by ≥25 sampling sites. No single ecoregion from the midwestern or western plains had 25 sites with paired samples, so 2 ecoregions, the Eastern Corn Belt Plains and Central Corn Belt Plains, were combined into a single ecoregion (Fig. 1). For each ecoregion, 2 data sets were constructed from the paired samples. One data set consisted of the samples from hard substrates, and the other consisted of the samples

from soft substrates. Both data sets had the same number of samples, and the samples from each site were collected at the same time.

Nonparametric sign tests were used to determine whether the structural attributes of algal assemblages differed between hard and soft substrates. The following attributes were compared: total algal biovolume, biovolume of selected algal groups, species diversity, and proportion of motile diatoms. Algal biovolumes were determined using the procedure described in Charles et al. (2002). The total number of algal species in a sample (species richness) and the Shannon–Wiener diversity index based on relative abundances of all algal taxa in a sample were used as measures of community diversity. Diversity measures and proportion of motile diatoms in a diatom count are commonly used in algae-based water-quality assessments (e.g., Stevenson and Bahls 1999).

Indicator species analysis (Dufrêne and Legendre 1997) was used to identify species associated with 4 types of substrate (rock, wood, sand, and silt) in the ecoregion data sets. This method identifies species that have the highest specificity (mean relative abundance) and fidelity (frequency of occurrence) to a certain habitat. The method does not require equal numbers of samples from each habitat, so all available NAWQA samples were used. The number of samples in the ecoregion-scale data sets ranged from 89 (Southeastern Plains) to 291 (Eastern and Central Corn Belt Plains). Indicator species analysis was carried out with PC-ORD/4 (MjM Software, Gleneden Beach, Oregon).

Pairwise comparisons among a series of ordinations were used to determine whether algal assemblages from hard and soft substrates responded to the same environmental gradients within each ecoregion. Non-Metric Multidimensional Scaling (NMS) ordinations were carried out using PC-ORD/4 on paired data sets representing hard and soft substrates in the 9 ecoregions. The relative abundances of all algal taxa were square-root transformed. A 2-dimensional solution was chosen to facilitate graphic representation of the ordination results. The distance measure was the Bray–Curtis dissimilarity coefficient. Procrustes Analysis (Gower 1971) with a permutation test developed by Peres-Neto and Jackson (2001) was used to compare ordinations of samples from soft and hard substrates in each ecoregion (PROTEST software, available from http://www.zoo.utoronto.ca/jackson/pro1.html). This analysis compares pairs of ordinations by applying a rotational-fit algorithm to find the optimal match between corresponding observations in the ordination results. The degree of correspondence between 2 ordinations is estimated by the m_{12} statistic, with lower numbers indicating higher degrees of correspondence.

Relationships between water-quality parameters and algal assemblages from different substrates in the ecoregion data sets were assessed using the following approach. A series of Canonical Correspondence Analyses (CCAs) were run with one environmental variable at a time for each of the 18 data sets (9 ecoregions, 2 substrate types/ecoregion). In 7 ecoregions, the hard substrate was rock; in 2, it was submerged wood. TP, TN, conductivity, and pH were used as constraining variables in the CCAs. These variables were chosen because they represent important water-quality characteristics, and they influence algal assemblage composition on an ecoregion scale (Potapova and Charles 2002). The strengths of relationships between algal assemblages and environmental variables were assessed using the ratios of the 1st and 2nd eigenvalues (λ_1/λ_2). This ratio measures the strength of the constraining variable with respect to the 1st unconstrained gradient in the assemblage composition data. Large numbers indicate strong responses of algal assemblages to the environmental variable (ter Braak and Prentice 1988). The strength of relationship is considered very high if $\lambda_1/\lambda_2 > 1$, moderately high if $0.5 < \lambda_1/\lambda_2 < 1$, and weak if $\lambda_1/\lambda_2 < 0.5$. Each CCA was run first with diatom data only, and second with all algal taxa. The total number of CCAs was 144 (9 ecoregions $\times$ 2 substrates $\times$ 4 constraining variables $\times$ 2 species sets). CCAs were carried out with CANOCO (version 4.5, Microcomputer Power, Ithaca, New York; ter Braak and Šmilauer 2002). Non-parametric Friedman tests were used to compare λ_1/λ_2 ratios among groups of CCAs.

Inference models were used to assess the utility of samples from soft or hard substrates for water-quality assessment. These models were constructed for several ecoregion data sets that showed strong relationships with chemistry parameters (i.e., with λ_1/λ_2 ratio >0.75 in the CCA). A ratio >0.5 is commonly considered an indication that the relationship between assem-

TABLE 1. Comparisons of median values of characteristics of algal assemblages from hard (1st value) and soft (2nd value) substrates in the National Water-Quality Assessment data set. * = $p < 0.05$, ** = $p < 0.005$.

Ecoregion (no. sample pairs)	Biovolume (mm^3/dm^2)				Species richness	Shannon–Wiener diversity index	% motile diatoms
	All algae	Diatoms	Green algae	Cyanobacteria			
Sierra Nevada (39)	53/80	36/64	0.50/0.11	2/0.6	46/55**	2.8/3.9**	17/28
Wasatch and Uinta Mountains (30)	73/173	33/109*	0.97/0.00*	2.2/1.6	31/43**	2.5/3.1**	29/30
Southern Rockies (61)	37/83*	18/69**	1.76/0.72	1.6/1.6	43/51**	3.0/3.6**	27/33
Piedmont (39)	33/128*	15/110**	0.11/0.11	1.6/0.9	45/58**	2.9/3.7**	15/56**
East and Central Corn Belt Plains[a] (33)	88/209	46/157	0.62/0.03	4.2/3.9	55/55	4.3/4.1	56/56
Northeastern Highlands (31)	52/182*	35/166**	6.31/3.65	3.7/1.1	40/48	2.2/3.4**	8/15**
Northeastern Coastal Zone (61)	12/86**	7/77**	0.02/0.06	0.2/0.1	37/52**	3.1/4.1**	10/18**
Southeastern Plains[a] (29)	14/69**	10/61**	0.06/0.11	0.4/2*	60/62	3.2/3.2	35/48**
Ridge and Valley (71)	69/193*	37/129**	0.19/0.00	2.5/0.8*	38/49**	3.2/4.2**	31/47**
All NAWQA sites (1048)	43/128**	21/104**	0.37/0.16**	1.3/1.0**	43/52**	3.8/3.9**	30/41**

[a] Hard substrate was rock or submerged wood

blage composition and environmental parameter is strong enough to justify inference models (Dixit et al. 1991). Several inference techniques were tested, including weighted-averaging (WA) with classical or inverse deshrinking with and without tolerance downweighting, WA partial least squares, maximum likelihood, and modern analogue. All models were carried out with the computer program C[2] (Juggins 2003) using square-root-transformed species relative abundance data. For each combination of ecoregion and environmental variable, 2 models were based on calibration data sets representing hard and soft substrates. A third model was based on a combined data set consisting of samples collected from both types of substrate. Two tests were used to determine if the models in each pair differed significantly in their performance. First, root mean square errors of prediction (RMSEP) jackknifed values were compared using an *F*-test on the ratio of the residual variances. Second, the significance of the differences of correlation coefficients (*R*-jackknifed) between observed and inferred values of the environmental parameter were tested using the *t* statistic. The models developed from each type of substrate were tested using the additional (unpaired) NAWQA samples collected from various types of substrates at the sites that were not used in model development.

Results

Differences in algal assemblages on hard and soft substrates

At the national scale, algal assemblages from hard substrates had significantly lower species richness, diversity, % motile diatoms, diatom biovolume, and total algal biovolume than assemblages from soft sediments (Table 1). The biovolumes of green algae and cyanobacteria were significantly higher on hard substrates. The same patterns were observed within some ecoregions (Table 1). No significant differences in any attribute were found in the East and Central Corn Belt Plains ecoregion.

Indicator species analysis showed that common algal taxa were not restricted to single substrates. Indicator values can vary from 0% for a taxon that has the same occurrence and abundance in all groups of samples to 100% for a taxon that is confined to one group of samples. In our analysis, indicator values rarely exceeded 60% (Table 2). However, some taxa had higher indicator values for hard substrates than for soft

TABLE 2. The 3 algal taxa with the highest indicator values (in parentheses) for hard and soft substrates in each ecoregion and substrate type. When the type of soft substrate was not listed, taxa are shown under both sand and silt. – indicates substrate not sampled.

	Hard substrate		Soft substrate	
Ecoregion	Rocks	Wood	Sand	Silt
Sierra Nevada	*Cocconeis placentula* var. *lineata* (53)	*Nitzschia palea* (56)	*Sellaphora pupula* (59)	
	Cocconeis pediculus (51)	*Navicula symmetrica* (53)	*Nitzschia frustulum* (59)	
	Rhoicosphenia abbreviata (48)	*Navicula minima* (52)	*Planothidium rostratum* (51)	
Wasatch and Uinta Mountains	*Chantransia* stage of undetermined red algae (45)	*Nitzschia amphibia* (56)	*Encyonema minutum* (47)	*Cocconeis placentula* var. *lineata* (48)
	Undetermined coccoid cyanobacteria (40)	*Rhiocosphenia abbreviata* (52)	*Achnanthidium minutissimum* (46)	*Navicula reichardtiana* (43)
	Calothrix parietina (30)	*Navicula lanceolata* (47)	*Staurosira construens* (41)	*Diatoma vulgaris* (42)
Southern Rockies	*Cymbella affinis* (34)	*Navicula subminuscula* (86)	*Fragilaria vaucheriae* (42)	*Planothidium lanceolatum* (40)
	Nitzschia dissipata (27)	*Navicula veneta* (53)	*Achnanthidium minutissimum* (37)	*Navicula gregaria* (28)
	Reimeria sinuata (27)	*Navicula symmetrica* (51)	*Sellaphora pupula* (32)	*Cocconeis placentula* var. *euglypta* (26)
Piedmont	*Chantransia* stage of undetermined red algae (62)	*Frustulia crassinervia* (48)	*Navicula cryptocephala* (72)	
	Encyonema minutum (50)	Undetermined pseudanabaenaceae (44)	*Geissleria decussis* (68)	
	Achnanthidium rivulare (49)	*Homoeothrix* sp. (39)	*Nitzschia palea* (63)	
Eastern and Central Corn Belt Plains	*Amphora pediculus* (54)	*Navicula subminuscula* (53)	*Gomphonema olivaceum* (45)	*Nitzschia palea* var. *debilis* (49)
	Chantransia stage of undetermined red algae (38)	*Simonsenia delongei* (44)	*Navicula gregaria* (44)	*Nitzschia palea* (49)
	Rhiocosphenia abbreviata (38)	*Nitzchia amphibia* (41)	*Navicula lanceolata* (41)	*Nitzchia acicularis* (46)
Northeastern Highlands	*Cocconeis placentula* var. *lineata* (36)	–	*Encyonema minutum* (76)	
	Rhoicosphenia abbreviata (31)	–	*Achnanthidium minutissimum* (75)	
	Undetermined Pseudanabaenaceae (28)	–	*Navicula cryptocephala* (70)	
Northeastern Coastal Zone	*Cocconeis placentula* var. *lineata* (51)	–	*Nitzschia palea* (82)	–
	Chantransia stage of undetermined red algae (51)	–	*Planothidium lanceolatum* (72)	–
	Nitzschia amphibia (48)	–	*Psammothidium bioretii* (67)	–
Southeastern Plains	–	*Gomphonema parvulum* (75)	*Nitzschia palea* (66)	–
	–	*Frustulia crassinervia* (71)	*Sellaphora pupula* (58)	–

TABLE 2. Continued.

Ecoregion	Hard substrate		Soft substrate	
	Rocks	Wood	Sand	Silt
	–	*Navicula cryptocephala* (47)	*Planothidium rostratum* (53)	–
Ridge and Valley	*Reimeria sinuata* (47)	–	*Nitzschia palea* (64)	
	Rhoicosphenia abbreviata (44)	–	*Achnanthidium minutissimum* (62)	
	Cocconeis pediculus (43)	–	*Nitzschia dissipata* (60)	

substrates (e.g., *Amphora pediculus*, *Rhoicosphenia abbreviata*, *Reimeria sinuata*, *Cymbella* spp., *Gomphonema* spp., and *Cocconeis* spp.). Most of these diatoms attach to substrates by stalks or mucilage. Motile diatoms, predominantly *Navicula* spp. and *Nitzschia* spp., were associated with soft sediments and with submerged wood, which was sampled mostly in low-gradient rivers. *Achnanthidium minutissimum*, commonly considered a disturbance-tolerant species (Stevenson and Bahls 1999), often was associated with soft sediments in the rivers of mountainous regions, but was associated with hard substrates in some regions with lower relief. For most nondiatom taxa, indicator values were <30%. Exceptions were some filamentous cyanobacteria (e.g., *Calothrix parietina*, undetermined Pseudanabaenaceae) and red algae that had higher indicator values for hard substrates than for soft substrates (Table 2).

Comparison of ordinations based on sets of samples from hard and soft substrates

Procrustes Analysis showed that NMS ordinations of samples from hard substrates were similar to ordinations of samples from soft substrates in all 9 ecoregions. The m_{12} statistic ranged from 0.36 to 0.77, with all p-values <0.001, indicating significant concordance between sample positions in ordination space. Overlays of environmental variables on ordinations showed that ordinations revealed similar environmental gradients underlying the structure of algal assemblages. Ordinations of soft and hard substrate samples were most similar (m_{12} = 0.36) for the Northeastern Coastal Zone ecoregion and least similar (m_{12} = 0.77) for the Wasatch and Uinta Mountains ecoregion (Fig. 2). In both ecoregions, ordinations of samples from soft and hard substrates were associated with similar major environmental gradients. In the Northeastern Coastal Zone, the major gradient in algal species composition corresponded to the environmental gradient from acidic, nutrient-poor rivers at high elevations to alkaline, nutrient-rich rivers at low elevations. In the Wasatch and Uinta Mountains, the major gradient in algal species composition corresponded to the environmental gradient from high-conductivity, nutrient-rich rivers at high elevations to low-conductivity, nutrient-poor rivers at low elevations.

Relationships between water-quality variables and algal assemblages

The strengths of relationships between algal assemblages and environmental variables (λ_1/λ_2 ratios) did not vary in a consistent manner between hard and soft substrates (Fig. 3). λ_1/λ_2 ratios from hard and soft substrates were not significantly different (Friedman's test, $p > 0.05$). Differences in relationship strengths were greater among data sets representing different ecoregions than among those representing different substrate types. Differences in relationship strengths attributed to ecoregions and environmental variables were significant (Friedman's test, $p < 0.05$). λ_1/λ_2 ratios were significantly higher for conductivity (0.61) than for TP (0.53), TN (0.43), and pH (0.45) ($p < 0.05$). Relationships with water chemistry generally were stronger for diatoms (λ_1/λ_2 ratio = 0.55) than other algal taxa (λ_1/λ_2 ratio = 0.48).

Applicability of inference models to samples collected from hard and soft substrates

Inference models were constructed for the 6 combinations of ecoregion and environmental

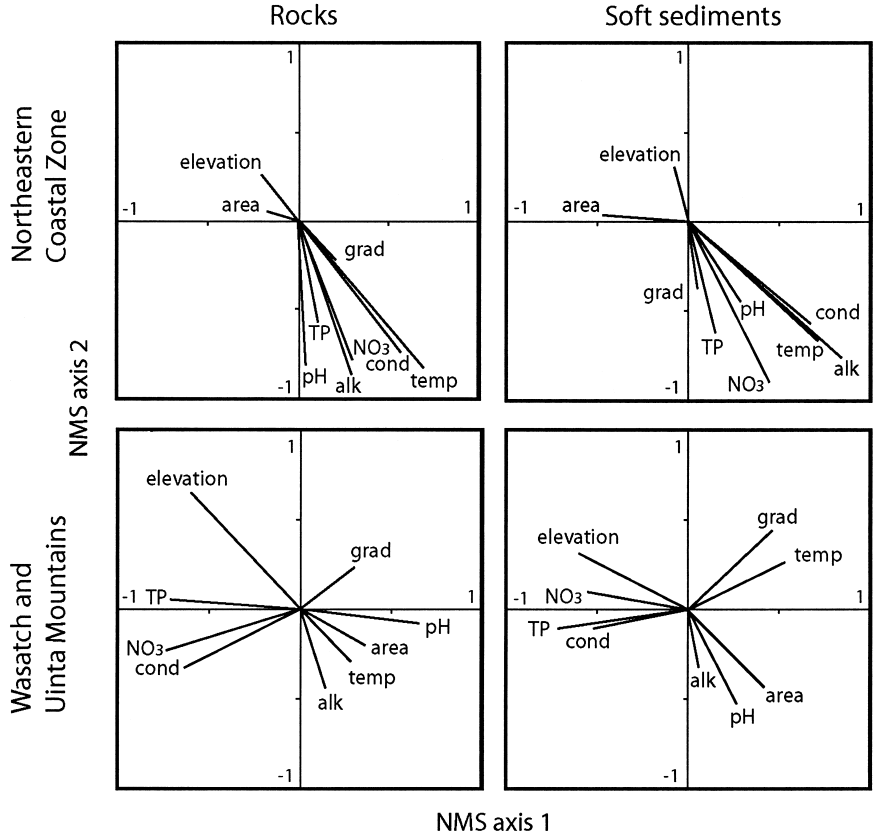

FIG. 2. Overlays of environmental variables on the Non-Metric Multidimensional Scaling (NMS) ordination plots after Procrustes rotation, showing similarity between ordinations of samples collected from rocks and soft sediments in 2 ecoregions. alk = alkalinity, area = drainage area, cond = conductivity, grad = channel gradient, temp = mean annual air temperature, TP = total P.

variables (Table 3), where CCAs for both hard and soft substrate data sets indicated a strong response to the environmental parameter ($\lambda_1/\lambda_2 > 0.75$). Only diatom data sets met this requirement. The inverse deshrinking WA models without tolerance downweighting usually had the lowest RMSEP, and only those models are reported here. Performance of models based on soft and hard substrate data sets (Table 3) did not differ significantly ($p > 0.05$ in all cases). When samples from hard and soft substrates were combined into a single data set for each ecoregion, the predictive power of the inference models did not improve significantly, but correlation coefficients between observed and inferred values of environmental variables increased in comparison with models based on single-substrate data sets (Table 3). The only exception was the conductivity model for the Southeastern Plains data set, which improved significantly after soft sediment and submerged wood data sets were combined. The models also were applied to unpaired test samples collected from various substrates at river sites in the same ecoregions. The quality of prediction estimated by RMSEP did not change significantly ($p > 0.05$ in all cases) when models constructed on the basis of one substrate were applied to samples from other substrates (Table 3).

Discussion

Substrate effects on community structure at national and ecoregion scales

Algal biovolume, diversity, and species composition differed significantly between hard and soft substrates at the national scale, and often,

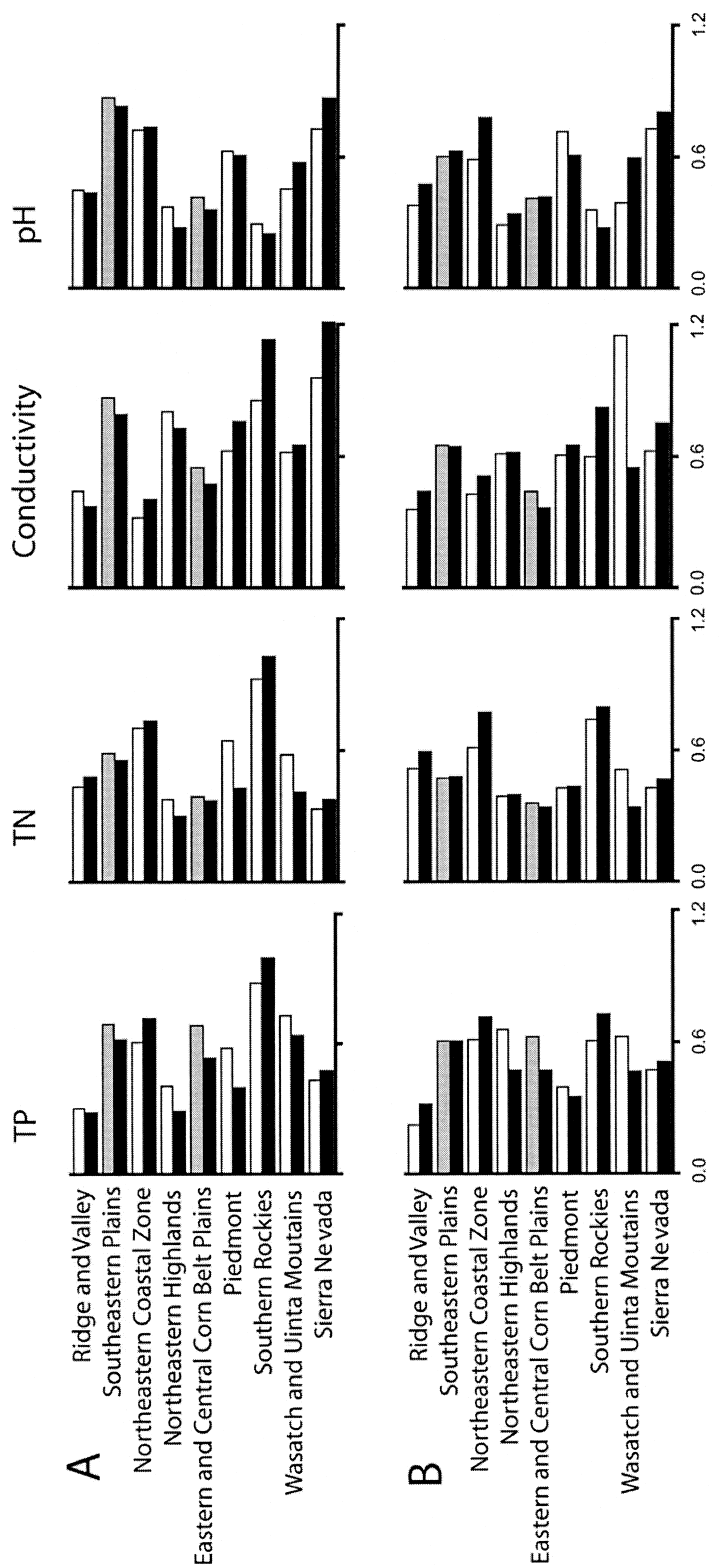

FIG. 3. Strengths of relationships between algal assemblages and total P (TP), total N (TN), conductivity, and pH expressed as the ratio of the 1st to the 2nd eigenvalues (λ_1/λ_2) in canonical correlation analyses carried out with single constraining variables. A.—Diatom data sets. B.—All algal taxa data sets.

TABLE 3. Comparisons of diatom weighted-averaging inference model performance for 3 data sets (hard substrate, soft substrate, both) in 3 ecoregions (Sierra Nevada, $n = 28$; Southern Rockies, $n = 44$; and Southeastern Plains, $n = 22$). Models developed from each data set were applied to unpaired test samples from different substrates. Values are correlation coefficients between observed and jackknifed inferred values of environmental parameters (*R*-jack) and root mean square errors of prediction (RMSEP). Hard substrates were rocks in the Sierra Nevada and Southern Rockies and submerged wood in the Southeastern Plains. * indicates a model developed from both substrates with performance significantly ($p < 0.05$) improved compared to models developed from single-substrate calibration data sets.

	Hard substrate		Soft substrate		Both	
Ecoregion/environmental variable	*R*-jack	RMSEP	*R*-jack	RMSEP	*R*-jack	RMSEP
Sierra Nevada/log conductivity (μS/cm)	0.90	0.22	0.92	0.19	0.92	0.19
Applied to 23 epidendric[a] samples		0.44		0.43		0.42
Applied to 10 epilithic samples		0.31		0.29		0.30
Applied to 18 soft-sediment samples		0.50		0.55		0.52
Southern Rockies/log conductivity (μS/cm)	0.77	0.31	0.77	0.32	0.82	0.29
Applied to 8 epidendric samples		0.16		0.10		0.12
Applied to 7 epilithic samples		0.63		0.41		0.52
Applied to 8 soft-sediment samples		0.21		0.09		0.07
Southeastern Plains/log conductivity (μS/cm)	0.77	0.18	0.72	0.19	0.91*	0.12*
Applied to 24 epidendric samples		0.38		0.39		0.39
Applied to 13 soft-sediment samples		0.24		0.24		0.24
Southeastern Plains/pH	0.69	0.39	0.72	0.37	0.84	0.29
Applied to 23 epidendric samples		0.49		0.43		0.46
Applied to 22 soft-sediment samples		0.58		0.57		0.57
Southern Rockies/log total P (μg/L)	0.76	0.41	0.81	0.38	0.82	0.35
Applied to 11 epidendric samples		0.62		0.30		0.29
Applied to 2 epilithic samples		0.28		0.57		0.61
Applied to 11 soft-sediment samples		0.36		0.11		0.20
Southern Rockies/log total N (μg/L)	0.87	0.28	0.86	0.31	0.88	0.27
Applied to 11 epidendric samples		0.44		0.43		0.49
Applied to 2 epilithic samples		0.26		0.28		0.29
Applied to 11 soft-sediment samples		0.50		0.34		0.45

[a] Collected from submerged wood

but not always, at the ecoregion scale. The power of statistical tests depends on the magnitude of the effect and the number of observations. An existing effect may not be detected if it is subtle and the number of observations is too small. Thus, we see 2 possible reasons why between-substrate differences were not always detected at the ecoregion scale. First, algal assemblages were more similar between soft and hard substrates in some ecoregions, such as the Central and Eastern Corn Belt, than in others. Second, our statistical tests were less powerful when applied to the ecoregion data sets than when applied to the national data set because ecoregion data sets consisted of fewer samples than the national data set.

Between-substrate differences in structural attributes of algal assemblages have been found most often in studies of single water bodies (Tuchman and Stevenson 1980) or in experimental settings (Burkholder 1996). In small-scale studies, substrate type is typically one of the major factors determining variability of algal assemblages. The influence of substrate usually is more difficult to detect in large-scale, coarse-resolution studies, such as ours, when the roles of other factors, such as between-stream differences in hydrology, physical habitat, and chem-

istry become more important than the role of substrate (Soininen and Eloranta 2004).

Substrate effects on relationships between assemblage composition and water quality

Algal assemblages on hard and soft substrates were associated with similar environmental gradients within ecoregions, despite differences in species composition (NMS). The strengths of relationships between algal assemblage composition and water chemistry parameters did not differ between hard and soft substrates (CCA). These results were unexpected because algae on fine-grained sediments are more influenced by sediment-bound chemicals (Wetzel 1983, Burkholder and Cuker 1991) than epilithic algae (Kelly et al. 1998), and are affected only weakly by water column chemistry. We attribute the similarities in the relationships of assemblages on hard and soft substrates to water chemistry parameters to the overriding effects of chemistry and other (unquantified) factors that masked the influence of substrate. Our conclusion is supported by Jüttner et al. (1996) and Rothfritz et al. (1997) who found that epiphytic and epilithic diatom assemblages in Nepalese streams varied similarly along water chemistry gradients because substrate influence was negligible compared to these factors.

Suitability of substrates for developing water-quality assessment techniques

The inference models that were developed from data sets representing different substrates did not differ significantly in their ability to infer water chemistry. Pan et al. (1996) compared performance of inference models based on diatom assemblages from erosional and depositional habitats in Appalachian streams. The predictive powers of their WA inference models based on pH and TP were approximately the same for both habitats. Similarity in predictive power of models based on calibration sets of samples collected from hard or soft substrates indicates that both are equally useful for water-quality assessment. Other techniques based on species autecologies, such as metrics or indices, are, in fact, simplified inference models (Potapova et al. 2004). Therefore, they should provide similar accuracy of water-quality assessment regardless of whether they were developed from samples collected from hard or soft substrates.

Transferability of water-quality assessment techniques between substrates

Algal assemblages differed between substrates, so we recommend using samples from similar substrates to compare attributes such as diversity, abundance of taxa with specific growth habits, or biovolume of major taxonomic groups. This recommendation concurs with Rosen (1995), Lowe and Pan (1996), and Kelly et al. (1998).

Inference models constructed using samples from one substrate inferred water chemistry equally well when applied to test samples from other substrates. Thus, models developed from hard and soft substrates appeared interchangeable. This result also suggests that only one sample has to be collected at each site for water-quality assessment surveys, and that sample can come from whatever substrate is available. This suggestion is based on the performances of only 18 models from 3 ecoregions, but is supported by other studies in which values of trophic and saprobic diatom indices did not differ whether they were derived from epilithon, epipelon, or epiphyton (Rott et al. 1998, Kitner and Poulíčková 2003). In contrast, values of some diatom indices did vary with substrate type in Finnish rivers. Indices were highest, indicating best water quality, for samples collected from water plants, slightly lower for samples from stones, and lowest for soft-sediment samples (Eloranta and Andersson 1998). However, the statistical significance of these differences was not established. Kelly et al. (1998) cited several studies in which slight differences were found between values of diatom indices calculated for epilithic and epiphytic assemblages. They recommended using rocks as a standard natural substrate for sampling algae. We agree that sampling standard substrates is a desirable way to eliminate the possible influence of substrate, and that standard substrates are necessary in small-scale studies carried out within single water bodies or small watersheds. However, when river surveys encompass diverse landscapes and different river types, a single preferred substrate type may not be available at all sites. In such cases, resources should be invested in collecting single samples from as many sites as possible, rather

than in sampling multiple substrates from fewer sites. The choice of substrate in surveys at the regional scale should not affect accuracy of water-quality assessments based on inference models or other autecological indices and metrics. However, if non-autecological structural attributes are used in assessments, algal samples should always be collected from a single substrate.

Acknowledgements

This paper was produced as part of a cooperative research agreement with the US Geological Survey NAWQA program. This manuscript is submitted for publication with the understanding that the US Government is authorized to reproduce and distribute reprints for governmental purposes. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing official policies, either expressed or implied, of the US Government. We are grateful to Carol Couch, Steve Moulton, and Peter Ruhl for their help in providing environmental data, to all NAWQA study unit biologists who collected samples, to Frank Acker, Robert Albrecht, Loren Bahls, Todd Clason, William Cody, Erik Jourdain, Michael Kachur, Kalina Manoylov, Lont Marr, Eduardo Morales, Sean Morris, Lara Panayotoff, Charles Reimer, Noma Ann Roberts, Tom Smith, Jackie White-Reimer, and Diane Winter for analyzing the algal samples, and to Kathleen Sprouffske for managing Phycology Section databases. We thank Steve Moulton, Daren Carlisle, 2 anonymous referees, and Scott Larned for helpful comments on the manuscript.

Literature Cited

BARBOUR, M. T., J. GERRITSEN, B. D. SNYDER, AND J. B. STRIBLING. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish. 2nd edition. EPA 841-B-99-002. Office of Water, US Environmental Protection Agency, Washington, DC.

BURKHOLDER, J. M. 1996. Interactions of benthic algae with their substrata. Pages 253–297 *in* R. J. Stevenson, M. L. Bothwell, and R. L. Lowe (editors). Algal ecology: freshwater benthic ecosystems. Academic Press, San Diego, California.

BURKHOLDER, J. M., AND B. E. CUKER. 1991. Response of periphyton communities to clay and phosphate loading in a shallow reservoir. Journal of Phycology 27:373–384.

CHARLES, D. F., C. KNOWLES, AND R. DAVIS. 2002. Protocols for the analysis of algal samples collected as part of the U.S. Geological Survey National Water-Quality Assessment Program. Patrick Center for Environmental Research Report No. 02–06. Patrick Center for Environmental Research, The Academy of Natural Sciences, Philadelphia, Pennsylvania. (Available from: http://diatom.acnatsci.org/nawqa/)

DIXIT, S. S., A. S. DIXIT, AND J. P. SMOL. 1991. Multivariate environmental inferences based on diatom assemblages from Sudbury (Canada) lakes. Freshwater Biology 26:251–266.

DUFRÊNE, M., AND P. LEGENDRE. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs 67:345–366.

ELORANTA, P., AND K. ANDERSSON. 1998. Diatom indices in water quality monitoring of some South-Finnish rivers. Verhandlungen der Internationalen Vereinigung für theoretische und angewandte Limnologie 26:1213–1215.

GOWER, J. C. 1971. Statistical methods of comparing different multivariate analyses of the same data. Pages 138–149 *in* F. R. Hodson, D. G. Kendall, and P. Tautu (editors). Mathematics in archaeological and historical sciences. Edinburgh University Press, Edinburgh, UK.

GURTZ, M. E. 1993. Design of biological components of the National Water-Quality Assessment (NAWQA) Program. Pages 323–354 *in* S. L. Loeb and A. Spacie (editors). Biological monitoring of aquatic systems. Lewis Publishers, Boca Raton, Florida.

HUGHES, R. M., AND D. P. LARSEN. 1988. Ecoregions: an approach to surface water protection. Journal of the Water Pollution Control Federation 60:486–493.

JUGGINS, S. 2003. C² user guide. Software for ecological and palaeoecological data analysis and visualisation. University of New Castle, New Castle upon Tyne, UK.

JÜTTNER, I., H. ROTHFRITZ, AND S. J. ORMEROD. 1996. Diatoms as indicators of river quality in the Nepalese Middle Hills with consideration of the effects of habitat-specific sampling. Freshwater Biology 36:475–486.

KELLY, M. G., A. CAZAUBON, E. CORING, A. DELL'UOMO, L. ECTOR, B. GOLDSMITH, H. GUASCH, J. HÜRLIMANN, A. JARLMAN, B. KAWECKA, J. KWADRANS, R. LAUGASTE, E.-A. LINDSTRØM, M. LEITAO, P. MARVAN, J. PADISÁK, E. PIPP, J. PRYGIEL, E. ROTT, S. SABATER, H. VAN DAM, AND J. VIZINET. 1998. Recommendations for the routine sampling of diatoms for water quality

assessments in Europe. Journal of Applied Phycology 10:215–224.

KITNER, M., AND A. POULÍČKOVÁ. 2003. Littoral diatoms as indicators for the eutrophication of shallow lakes. Hydrobiologia 506–509:519–524.

LOWE, R. L., AND Y. PAN. 1996. Benthic algal communities as biological monitors. Pages 705–739 *in* R. J. Stevenson, M. L. Bothwell, and R. L. Lowe (editors). Algal ecology: freshwater benthic ecosystems. Academic Press, San Diego, California.

MOULTON, S. R., J. G. KENNEN, R. M. GOLDSTEIN, AND J. A. HAMBROOK. 2002. Revised protocols for sampling algal, invertebrate, and fish communities as part of the National Water-Quality Assessment Program. U.S. Geological Survey Open-File Report 02–150. US Geological Survey, Reston, Virginia. (Available from: http://water.usgs.gov/nawqa/protocols/OFR02–150)

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. 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.

PERES-NETO, P. R., AND D. A. JACKSON. 2001. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. Oecologia (Berlin) 129:169–178.

PORTER, S. D., T. F. CUFFNEY, M. E. GURTZ, AND M. R. MEADOR. 1993. Methods for collecting algal samples as part of the National Water-Quality Assessment Program. U.S. Geological Survey Open-File Report 93–409. US Geological Survey, Raleigh, North Carolina. (Available from: http://water.usgs.gov/nawqa/protocols/OFR-93–409)

POTAPOVA, M. G., AND D. F. CHARLES. 2002. Benthic diatoms in USA rivers: distributions along spatial and environmental gradients. Journal of Biogeography 29:167–187.

POTAPOVA, M. G., D. F. CHARLES, K. C. PONADER, AND D. M. WINTER. 2004. Quantifying species indicator values for trophic diatom indices: a comparison of approaches. Hydrobiologia 517:25–41.

ROSEN, B. H. 1995. Use of periphyton in the development of biocriteria. Pages 209–215 *in* W. S. Davis and T. P. Simon (editors). Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

ROTHFRITZ, H., I. JÜTTNER, A. M. SUREN, AND S. J. ORMEROD. 1997. Epiphytic and epilithic diatom communities along environmental gradients in the Nepalese Himalaya: implications for the assessment of biodiversity and water quality. Archiv für Hydrobiologie 138:465–482.

ROTT, E., H. C. DUTHIE, AND E. PIPP. 1998. Monitoring organic pollution and eutrophication in the Grand River, Ontario, by means of diatoms. Canadian Journal of Fisheries and Aquatic Sciences 55:1443–1453.

ROUND, F. E. 1981. The ecology of algae. Cambridge University Press, Cambridge, UK.

SABATER, S., S. V. GREGORY, AND J. R. SEDELL. 1998. Community dynamics and metabolism of benthic algae colonizing wood and rock substrata in a forest stream. Journal of Phycology 34:561–567.

SOININEN, J., AND P. ELORANTA. 2004. Seasonal persistence and stability of diatom communities in rivers: are there habitat specific differences? European Journal of Phycology 39:153–160.

STEVENSON, R. J., AND L. L. BAHLS. 1999. Periphyton protocols. Pages 6/1–6/22 *in* M. T. Barbour, J. Gerritsen, B. D. Snyder, and J. B. Stribling. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish. 2nd edition. EPA 841-B-99-002. Office of Water, US Environmental Protection Agency, Washington, DC.

STEVENSON, R. J., AND S. HASHIM. 1989. Variation in diatom community structure among habitats in sandy streams. Journal of Phycology 25:678–686.

TER BRAAK, C. J. F., AND I. C. PRENTICE. 1988. A theory of gradient analysis. Advances in Ecological Research 18:271–317.

TER BRAAK, C. J. F., AND P. ŠMILAUER. 2002. CANOCO reference manual and CanoDraw for Windows user's guide: software for canonical community ordination (version 4.5). Microcomputer Power, Ithaca, New York.

TUCHMAN, M. L., AND R. J. STEVENSON. 1980. A comparison of clay tile, sterilized rock and natural substrate diatom communities in a small stream in southeastern Michigan, USA. Hydrobiologia 75:873–879.

WETZEL, R. G. 1983. Attached algal-substrata interactions: fact or myth, and when and how? Pages 207–215 *in* R. G. Wetzel (editor). Periphyton of freshwater ecosystems. Dr W. Junk Publishers, The Hague, The Netherlands.

Received: 17 December 2003
Accepted: 13 December 2004