



Benthic diatoms in USA rivers: distributions along spatial and environmental gradients

Marina G Potapova* and Donald F Charles *Patrick Center for Environmental Research, The Academy of Natural Sciences, Benjamin Franklin Parkway, 1900 Philadelphia, PA, 19103 USA*

Abstract

Aim To study large-scale patterns of benthic diatom assemblages in rivers, to assess the relative importance of environmental and geographical factors affecting their composition, and to evaluate the implications of these patterns for the use of diatoms as indicators of water quality, particularly nutrient enrichment.

Location The United States Geological Survey (USGS) National Water Quality Assessment (NAWQA) Programme data set covers the conterminous United States.

Methods We employed gradient analysis to understand the floristic structure of the data set and to discover major ecological gradients underlying variation in species composition at different spatial scales (entire US, Omernik Level 1 and Level 2 ecoregions). We used variance partitioning to separate the effect of environmental and spatial characteristics.

Results At the national scale, three major ecological gradients were evident. The first was a complex 'downstream' gradient from fast-flowing, mostly oligotrophic, highland rivers to predominantly eutrophic rivers of high- and low-elevation plains. The second was a gradient integrating water mineral content and pH, which separated the soft and often more acidic waters of the humid eastern part of US from the alkaline waters of arid western regions. The third obvious gradient was related to latitudinal and altitudinal variation of temperature. Up to one-third of the total explainable variation in species data was attributed solely to geographical factors not correlated with measured environmental characteristics. We present several examples of species with complex patterns of spatial distribution.

Main conclusions Although environment consistently plays the most important role in structuring diatom assemblages in rivers, spatial factors also explain some variation in diatom distribution, especially at the continental scale. Most of the species that are confined to limited geographical areas are not yet described and await future taxonomic work. We show that selection of species which could serve as indicators of nutrient enrichment in US rivers is not a straightforward procedure. The existence of complex environmental gradients, and still poorly understood spatial patterns of species distribution, precludes attempts to develop uniform diatom-based metrics that would be applicable everywhere in the US. We advocate the development and calibration of metrics based on data sets collected from more limited geographical areas, and that include sites having relatively narrow ranges of environmental characteristics other than that which the metrics are designed to indicate.

Keywords

Benthic diatoms, rivers, multivariate analysis, ecoregions, ecological gradients, biogeography, water quality indicators.

*Correspondence: Patrick Center for Environmental Research, The Academy of Natural Sciences, 1900 Benjamin Franklin Parkway, Philadelphia, PA 19103, USA. E-mail: potapova@acnatsci.org

INTRODUCTION

Diatoms are a large and diverse group of single-celled algae. They are distributed throughout the world in nearly all types of aquatic systems and are one of the most important food resources in marine and freshwater ecosystems. Because there are many thousands of taxa with diverse ecological requirements, their siliceous remains are used extensively as environmental indicators in studies of climate change, acidic precipitation and water quality (Stoermer & Smol, 1999). Despite their ecological importance, practical usefulness, and previous study by taxonomists and ecologists, large-scale patterns of diatom taxa distributions and their underlying causes are largely unexplored. In this paper we examine national- and regional-scale patterns of benthic diatom assemblages in rivers across the continental United States and assess the relative importance of environmental and geographical factors affecting their composition. We also discuss the implications of those patterns for the use of diatoms as indicators of water quality, in particular, nutrient enrichment.

A key issue in understanding diatom distributions is knowing the extent to which they are constrained by geographical factors that limit species dispersal vs. the extent to which they are limited only by the ability of the species to grow under a specific combination of environmental factors. Is a species not in a particular habitat only because it was not dispersed there or because, though it may have reached the habitat, the environmental conditions were unsuitable for it to survive and compete? Conversely, is a species in a specific location only because it can compete better under those particular environmental conditions, or might it also be there simply because it reached a site and its potential competitors did not? Dispersal is influenced by many factors, including physical barriers (deserts, mountains), distance between suitable habitats (density of water bodies), transportability of cells, type and importance of transport mechanisms (e.g. wind, bird's feet), the size and distribution of original populations, and the amount and distribution of habitat suitable for colonization (Kristiansen, 1996). Ability of populations to compete and persist in a new local habitat depends on how well adapted they are to grow under its environmental conditions. Some of the factors most often found to be important for distribution of benthic river diatoms are water chemistry (particularly pH, ionic strength and nutrient concentrations), substrate, current velocity, light and grazing (Patrick & Reimer, 1966; Round, 1981; Stevenson *et al.*, 1996). Most of these factors depend strongly on climate, geology, topography, land-use and other landscape characteristics, and therefore are similar within ecological regions defined by these characteristics (Stevenson, 1997). Many diatom species are known to have world-wide distributions; others appear limited to certain climatic zones or geographical regions, or are endemic to particular water bodies. Kociolek & Spaulding (2000) argue that the proportion of geographically restricted diatom species is much higher than formerly thought, and because of that, the importance of geograph-

ical factors in explaining patterns in diatom flora has been underestimated.

Understanding the relation of geographical and environmental factors to diatom distributions is important to the process of developing diatom-based water quality indicators. The information value of indicators depends largely on how they are developed and calibrated, and in particular on how well the ecological characteristics of diatom taxa are quantified. In this regard, it is useful to know what factors limit distribution of species and how the distribution varies with spatial scale and ranges of environmental characteristics. We suppose that if the occurrence of a diatom is limited primarily by environmental characteristics, and not by geographical factors that limit or direct dispersion, then that diatom should be found in similar environmental conditions no matter where they occur. If, however, geographical and chance dispersal factors have the greatest influence on species distributions, then it is more likely that a species will evolve and adapt better to its local environment and that ecological characteristics of the taxon will vary more widely. A key aspect of quantifying ecological data on indicator taxa is determining the geographical extent of sites from which to collect data and whether and how to limit the range of environmental characteristics of those sites. Any study designed to address this issue requires a large set of taxonomically consistent diatom data coupled with environmental and geographical information.

Substantial information on benthic diatoms of United States rivers has been accumulated over the last half of the twentieth century. In 1950–70s, scientists at the Academy of Natural Sciences in Philadelphia (ANS), led by Dr Ruth Patrick, surveyed many rivers in the eastern U.S. These surveys focused on the use of diatoms as water quality indicators (Patrick & Roberts, 1979; Patrick, 1996), but also led to the discovery of many new species (Patrick, 1945, 1946, 1958, 1959; Wallace, 1960; Hohn & Hellerman, 1963; Patrick & Reimer, 1966, 1975; Reimer, 1966). The works of Camburn *et al.* (1978) and Lowe & Kociolek (1984) brought new additions to the river diatom flora of this region, while studies of Pan, Stevenson and co-authors (Pan *et al.*, 1996, 1999, 2000) concentrated on the application of diatoms as indicators of water quality. Lowe and Stevenson, with colleagues, carried out most of their pioneering studies of the functional aspects of lotic algal communities in mid-western states (Kline & Lowe, 1975; Stevenson & Pryfogle, 1975; Jackson & Lowe, 1978; Millie & Lowe, 1981; Stevenson, 1983, 1990; Keithan & Lowe, 1985; Krejci & Lowe, 1987; Stevenson & Hashim, 1989; Stevenson & Peterson, 1989, 1991). Bahls (1982, 1983) studied both floristic and ecological (Bahls *et al.*, 1984) aspects of diatoms in Montana rivers. These studies, together with other numerous works on river diatoms (e.g. Hansmann, 1973; Benson & Rushforth, 1975; Lawson & Rushforth, 1975; Clark & Rushforth, 1977; Czarnecki & Blinn, 1978; Kennedy & Pfister, 1984; Rushforth & Merkley, 1988; Edlund, 1994), constitute an immense amount of information on the composition of benthic diatom communities in rivers across United States. Still, an

attempt to summarize these data to capture the broad-scale patterns of benthic diatoms would be a difficult mission, the main obstacle being differences in taxonomic concepts and sampling designs used by the investigators.

We analysed data from the United States Geological Survey's (USGS) National Water Quality Assessment (NAWQA) Programme. The NAWQA programme collects physical, chemical and biological data to monitor water quality in rivers across the U.S., and to analyse factors affecting it (Gurtz, 1993; <http://water.usgs.gov/nawqa/>). Eutrophication is recognized as a major water quality problem and special attention is given to assessment of nutrients. Fish, benthic invertebrates and algal communities are studied to help evaluate water quality and ecological conditions. The NAWQA programme used uniform procedures to collect algal samples throughout the U.S. These samples were analysed in a consistent manner, and the resulting data provide a unique opportunity to study diatom species distributions at the continental scale. Our future goal is to use these data to develop metrics and indices for water quality assessment. Despite the apparently large body of accumulated knowledge on diatom ecology, it is clear that many indices developed for particular geographical areas and water systems do not work well in other areas (Eloranta & Kwandrans, 1996). The reasons for this failure are typically that there is limited overlap in species composition between the two areas, or that at least some ecological characteristics of key taxa vary between the regions. Thus, it is important to explore major patterns in species composition before trying to develop new indices or apply already existing indices to the new geographical areas.

Using the NAWQA data, we sought to address several questions: (1) Are there regions of the country with benthic diatom floras distinct from other areas? Where are these floras and what species do they include? (2) To what extent can diatom assemblage composition be explained by environmental characteristics alone? By spatial variables alone?

Which ecological factors are most influential and how does their importance vary among regions? (3) What are the implications of observed patterns of diatom distributions for the development of diatom-based metrics or indices for water quality assessment?

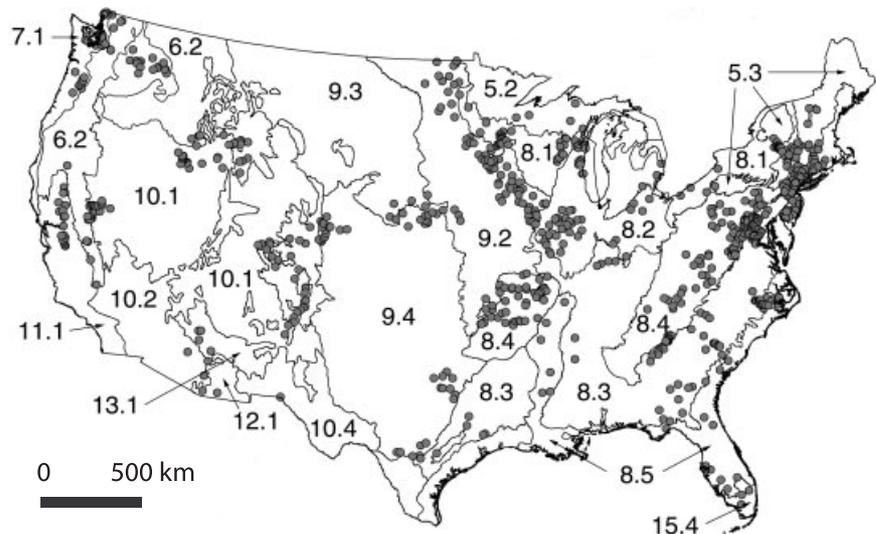
Our main approach to address these questions was, first, to detect major patterns in species composition of diatom assemblages and second, to find which ecological factors explained most of that variation. We used variance partitioning to separate the effect of environment and geographical factors on diatom assemblages (Borcard *et al.*, 1992). Hierarchical spatial classification of aquatic ecoregions proposed by Omernik (1995) served as the spatial framework of our analysis.

METHODS

Data collection

We used data for 2735 quantitative benthic samples collected from 1993 to 1998 at 897 sampling sites (Fig. 1). At some of the sites algal samples were taken only once, while at other sites they were taken once a year during two or three consecutive years (Gurtz, 1993). A full suite of environmental data was available for 247 sites only. These data included hydrology, water chemistry, physical watershed characteristics and land-use. The USGS personnel collected algal samples following standard procedures (Porter *et al.*, 1993). Quantitative samples were taken from two major types of habitat. 'Richest-targeted habitat' (RTH) samples were collected from rocks or wood in areas of relatively fast current (e.g. riffles or runs). 'Depositional-targeted habitat' (DTH) samples were collected from pools or other relatively slow-current locations. Depositional-targeted habitat samples were collected mainly from soft sediments, except in the Ozark Plateau, where they were collected from stones because there were no soft sediments at the sampling sites. Most often, both types of samples – RTH and DTH – were

Figure 1 Omernik's Level 2 Ecoregions of the continental United States and position of National Water Quality Assessment (NAWQA) sampling sites. Ecoregions where algal samples were taken: 5.2, Mixed Wood Shield; 5.3, Atlantic Highlands; 6.2, Western Cordillera; 7.1, Marine West Coast Forest; 8.1, Mixed Wood Plains; 8.2, Central Plains; 8.3, South-eastern Plains; 8.4, Ozark, Quachita, Appalachian Forests; 8.5, Mississippi Alluvial and South-eastern Coastal Plains; 9.2, Temperate Prairies; 9.3, West-Central Semi-Arid Prairies; 9.4, Southcentral Semi-Arid Prairies; 10.1, Western Interior Basin and Ranges; 10.2, Sonoran and Mohave Deserts; 10.4, Chihuahuan Desert; 11.1, Mediterranean California; 12.1, Western Sierra Madre Piedmont; 13.1, Upper Gila Mountains; 15.4, Everglades.



collected simultaneously at a sampling site. Sometimes, multiple samples of the same type were taken from a site in order to assess within-site variability. Algal samples were collected most often during low-flow conditions, usually in June–October.

Diatom slides and counts

Permanent diatom slides were prepared by oxidizing organic material in samples with nitric acid in a laboratory microwave oven and mounting cleaned diatoms in Naphrax. Diatom analysts at the Patrick Center of The Academy of Natural Sciences, Biology Department of University of Louisville, and independent contractors identified and counted diatoms. They counted 600 diatom valves on each slide, except for several samples where diatoms were too scarce. All slides used for diatom identification and enumeration were deposited in The Academy of Natural Sciences Diatom Herbarium.

Taxonomy

The diatom floras used for identification included Hustedt (1930a, b, 1959, 1961–66), Patrick & Reimer (1966, 1975), Krammer & Lange-Bertalot (1986, 1988, 1991a, b), Simonsen (1987), and Camburn *et al.* (1984–86), but regional floristic papers by Hohn & Hellefrman (1963), Lowe & Kociolek (1984), Kociolek *et al.* (1995), Camburn *et al.* (1978), Lowe (1972), and monographs by Lange-Bertalot & Moser (1994), Kociolek & Stoermer (1988, 1990, 1991), Kociolek & Kingston (1999), and Reichardt (1997) were also consulted. The eight analysts who identified and counted diatoms did not always follow the same taxonomic concepts. To mitigate possible problems this could cause in data analysis, we reviewed the taxonomy used by all analysts, and corrected most of the inconsistencies we found in the counts. The initial work on the taxonomic consistency in the NAWQA data set was carried out by J.C. Kingston.

Approximately 26% of the total number of diatom taxa found in 2735 samples could not be ascribed to any validly published diatom species. Some of these taxa clearly were yet undescribed species, while others were found in numbers insufficient to evaluate their taxonomic status. Analysts entered these 'unknown' taxa in the data base using temporary codes, as for example 'Cymbella sp. 1 ANS POTO'. These codes indicate location of the sample, initials of the analyst, or both, to facilitate future taxonomic work. In this paper we use names as they are stored in our data base. These undescribed taxa were documented by circling specimens on slides and making digital images. Our taxonomic concepts are documented in the ANS Algae Image Data base which can be accessed at <http://diatom.acnatsci.org>.

The main consequence of taxonomic differences among results produced by different analysts was that in a few instances, in order to combine counts into a single data set, we had to lump taxa. In two cases, we had to lump well established taxa into a single category. In the first case, we

combined together taxa that would be identified as *Achnanthes lanceolata* ssp. *rostrata* (Østrup) Lange-Bertalot, *A. lanceolata* ssp. *frequentissima* Lange-Bertalot, and *A. rostrata* ssp. *dubia* (Grunow) Lange-Bertalot, if using Krammer & Lange-Bertalot (1991b) as a reference. Analysts using Patrick & Reimer (1966) as a main taxonomic reference had identified those taxa as *A. lanceolata* var. *dubia* Grunow, although in fact in most cases it was *A. rostrata* Østrup [= *A. lanceolata* ssp. *rostrata* (Østrup) Lange-Bertalot]. As these diatoms were extremely common, recounting was not possible and all three taxa were lumped under the name '*A. rostrata*'. In the second case we had to lump together *A. biasolettiana* Grunow and *A. deflexa* Reimer because analysts had different and sometimes confusing concepts of these taxa. In other cases our lumping could be better described as an application of a broad species concept. For example, some analysts distinguished *Navicula gregaria* Donkin and *N. secreta* var. *apiculata* Patrick, while others used only one of those names for a group of taxa known as *N. gregaria* Donkin *sensu lato*. Another example is *Gomphonema pumilum* (Grunow) E. Reichardt et Lange-Bertalot. Several populations showed considerable variation in shape and could possibly represent separate species, but their taxonomic status is currently unclear and we preferred to lump them together. Lumping decreases resolution of the ecological analysis, because even closely related taxa can have different ecological preferences, and we tried to avoid doing it whenever possible.

Environmental data

United States Geological Survey's staff provided site and watershed information and water chemistry data. Geographical location information included latitude, longitude and ecoregion (accordingly to Omernik, 1995). This ecoregion classification of landscapes is based on climate, geology, soil, topography and potential vegetation. This classification is appealing because it accounts for the main ecological factors thought to directly or indirectly influence diatom distributions. Ecoregions are hierarchically arranged so that Level 2 Ecoregions are aggregated into Level 1 Ecoregions. In the text of this paper, the ecoregion rank, in brackets, follows the ecoregion name. Location of sampling sites within Level 2 Ecoregions is shown in Fig. 1. We used the following watershed characteristics: mean watershed elevation, mean watershed slope, mean annual air temperature (MAT), long-term water discharge (Q), watershed area, and percentage of major watershed land-use type (agricultural, urban, forest and range land). Substrate sampled was either rocks or submerged wood (branches of trees or woody debris) in the case of RTH samples, and rocks or soft sediment in the case of DTH samples. Water chemistry and temperature usually were measured at least once a month. Information on the analytical procedures used in the USGS NAWQA programme can be found at <http://water.usgs.gov/nawqa/>. We used the measurements of samples taken closest to the date when algae were collected. The following water chemistry parameters were used in our analyses: pH; conductivity; and

concentrations of Ca^{++} , Mg^{++} , Na^+ , K^+ , Cl^- , SO_4^- , total phosphorus (TP), total nitrogen (TN), dissolved nitrates plus nitrites ($\text{NO}_3\text{-N} + \text{NO}_2\text{-N}$), and dissolved ammonia ($\text{NH}_4\text{-N}$).

Data analysis

We used detrended correspondence analysis (DCA) to search for the main patterns in diatom assemblages, and to assess similarity among ecoregions in terms of assemblage composition. The DCA is an ordination technique that is useful for exploring assemblage patterns without making a priori assumptions about possible factors influencing assemblage structure. For DCA we used a data set consisting of all available 2735 samples representing 897 sites. It might be argued that by including more than one sample per site we biased the resulting ordination results toward more frequently sampled sites. However, we believe that considering the spatial and temporal variability of benthic communities in rivers, the larger number of samples better characterize species–environment relations for individual sites. Because we were interested only in species that live in benthic communities, we excluded all planktonic species from the analyses. Only the 433 taxa that reached at least 1% relative abundance in at least six samples were included in the 2735-sample data set.

We used canonical correspondence analysis (CCA) to relate assemblage structure to environmental variables and to explore how these relationships varied with geographical scale, and among ecoregions at the same scale. We began by constructing files of diatom count data for different geographical areas: (1) the entire country, (2) five Level 1 Ecoregions, and (3) five Level 2 Ecoregions. For constructing these data sets we used only 582 RTH samples from the 247 sites that had an accompanying complete set of environmental data, and excluded all replicate samples, so that only one sample represented each site/date combination. We also constructed a data set that included 247 RTH samples from the whole USA, but with only one sample per site. We included in our analyses all species that reached relative abundance of 1% in at least one sample from the corresponding data set.

We constructed separate blocks of environmental and spatial variables. The block of environmental variables included physical site characteristics, land-use (expressed as percentage of each of the four land-use types in a watershed), chemical variables, water temperature, and type of substrate that was sampled (rocks or wood). All physical variables were log-transformed, except those expressed as percentages. The four land-use variables and slope were square-root transformed. All chemical variables except pH were also log-transformed. The watershed area was excluded from the analyses because it was highly correlated with long-term discharge.

To create the block of spatial variables, we centred decimal latitude (X) and longitude (Y) for each sampling site on their means, and then calculated second- and third-degree monomials combining X and Y (Legendre & Legendre,

1998). A complete spatial set consisted of nine variables: X, Y, XY, X^2 , Y^2 , X^2Y , XY^2 , X^3 , Y^3 .

We ran preliminary CCAs separately with environmental and spatial sets to identify collinear variables (those with inflation factor higher than 20) and excluded them from final analyses (ter Braak, 1996).

To partition the variance of the species data explained by environmental and spatial factors, we used partial CCA as described by Borcard *et al.* (1992). This technique is applied to multivariate data when it is necessary to separate the effects of several sets of variables.

Species abundance data were log-transformed in all analyses. All ordinations were done using CANOCO version 4.1 (ter Braak & Šmilauer, 1998). CALIBRATE version 0.61 (Juggins & ter Braak, 1992) was used to calculate abundance weighted means when exploring latitudinal patterns of species distributions. Species distribution maps were created using the ArcView GIS program, version 3.2 (Environmental System Research Institute, Inc., CA, USA).

RESULTS

Environmental characteristics of study sites within ecoregions

The sites for which diatom and environmental data were available are located in ten Level 1 Ecoregions. Table 1 shows the numbers of sites and corresponding diatom samples for each ecoregion, as well as site elevation and slope, watershed land-use and selected water chemistry characteristics. Because Ecoregion 'Eastern Temperate Forest' (8), had a much larger number of sites/samples collected than the others, data are presented for five of its Level 2 Ecoregions.

There were obvious differences in site characteristics among ecoregions. These seem clear although the highly uneven distribution in number of sites and samples among ecoregions precludes rigorous statistical comparisons. The ecoregions with highest watershed slopes were 'North-western Forested Mountains' (6), 'Mediterranean California' (11), and 'North American Deserts' (10); those with the lowest were 'Tropical Wet Forest' (15), and 'Mississippi Alluvial and South-eastern Coastal Plain' (8.5). The highest proportions of agricultural land-use were in Ecoregions 'Eastern Temperate Forests' (8), 'Great Plains' (9), and 'Tropical Wet Forest' (15); the highest proportion of urban land-use is in Ecoregion 'Marine West Coast' (7). The median values for nutrients were highest in ecoregions with the largest percentage of agricultural and urban land-use. Lowest median pH values were recorded in south-eastern regions, while the highest were in the arid regions of the West.

Species composition of diatom assemblages

A total of 1548 diatom taxa was recorded in 2735 diatom samples. Of those taxa, eighty-seven are known as clearly preferring planktonic habitat and were excluded from further analyses.

Table 1 Selected chemical and physical parameters in National Water Quality Assessment (NAWQA) data sets containing only samples from sites that had an accompanying complete set of environmental data (first line: minimal, second line (in bold): median, third line: maximal value)

Level 1 Ecoregion Level 2 Ecoregion	Sites										Samples				
	Number of sites	Elevation (m a.s.l.)	Slope (%)	Air temperature (°C)	Land-use type (%)			Water temperature (°C)	Conductivity, $\mu\text{S cm}^{-1}$	pH	TN (mg L^{-1})	TP (mg L^{-1})			
					Forest	Range land	Agri-culture						Urban	Number of samples	
5 Northern Forests	10	334 443 678	0.14 0.76 10.26	2.4 4.6 7.7	15.0 78.4 99.3	0 4.0 54.0	0 0.3 1.7	0	26	10	26	5.8	<0.01		
6 Northwestern Forested Mountains	15	548 2628 3408	4.90 10.03 14.42	1.5 3.7 9.6	31.0 65.5 100	0 0.2 17.7	0 0.3 8.3	2	38	2	39	7.2	<0.01		
7 Marine West Coast Forest	8	40 123 1017	0.28 1.95 18.65	9.8 10.3 11.5	0 20.7 89.3	0 15.4 99.0	0 13.9 100	7	13	7	81	7.4	<0.01		
8 Eastern Temperate Forests	127	9 282 1073	0.04 1.17 8.09	4.3 11.1 22.7	0 34.3 97.9	0 42.1 98.6	0 2.2 100	2	306	2	20	5.6	<0.01		
8.1 Mixed Wood Plains	17	93 288 454	0.25 0.63 2.58	4.3 7.2 11.2	0.9 16.3 54.8	0 27.2 92.9	0.3 3.2 86.7	4	44	4	165	6.0	<0.01		
8.2 Central USA Plains	19	100 235 422	0.22 0.65 7.13	6.4 8.6 11.1	0 6.4 69.7	0 73.2 98.6	0.6 3.2 100	2	49	2	192	6.9	<0.01		
8.3 Southeastern USA Plains	35	12 153 402	0.07 0.89 1.94	8.3 15.4 19.8	0.02 29.6 93.0	0 41.6 92.8	0.01 6.9 100	4	86	4	20	5.6	<0.01		
8.4 Ozark, Ouachita-Appalachian Forests	45	130 497 1073	0.62 3.15 8.09	8.1 11.0 15.1	1.3 60.3 97.9	0 30.9 88.4	0 1.4 78.0	6	96	6	29	6.0	<0.01		
8.5 Mississippi Alluvial & Southeastern Coastal Plain	11	9 47 111	0.04 0.24 0.51	12.2 17.0 22.7	3.9 11.6 44.8	0 58.7 94.0	0.4 1.4 33.3	19	31	19	50	5.6	<0.01		
9 Great Plains	44	182 465 2606	0.23 0.82 6.58	2.7 8.3 20.0	0 7.1 77.8	0 40.5 99.7	0 1.1 95.9	11	99	11	122	7.1	<0.01		
10 North American Deserts	31	326 2087 3230	0.00 6.92 16.94	2.9 8.2 11.0	0 44.0 77.6	0 7.8 96.1	0 0.6 12.8	6	68	6	89	7.8	<0.01		

11 Mediterranean California	7	46 1275 1513	0.46 8.43 8.92	10.9 12.1 16.2	0 68.4 77.5	0 14.1 21.5	0 6.8 13.3	1.0 2.2 100	12	12	42	7.0 7.6 8.1	<0.2 0.3 2.1	<0.01 0.03 0.26
12 Southern Semi-Arid Highlands	1	1490	4.65	16.1	16.9	78.2	1.8	3.1	3	11	517	8.2	<0.1	<0.01
13 Temperate Sierras	1	2018	3.98	14.4	78.9	21.0	0	0.1	3	11	357	8.4	<0.1	<0.01
15 Tropical Wet Forests	3	3 5 5	0.01 0.02 0.03	22.8 23.0 23.4	0 0.4 11.8	0 17.0 29.4	0.01 65.1 95.8	0 0.3 2.4	14	28	160	6.9 7.2 7.4	0.8 1.3 4.1	<0.01 0.03 0.34
All sites/samples	247	3 378 3408	0.00 1.47 18.65	1.5 10 23.4	0 36 100	0 0 95	0 28.4 99.7	0 1.4 100	582	2	20	5.6 7.9 9.2	<0.1 1.0 14.6	<0.01 0.06 2.80

Of the 1461 non-planktonic diatom taxa, 1016 (70%) have been recorded from continents other than North America (Hustedt, 1930a, b, 1938–39, 1959, 1961–66; Hohn & Hellerman, 1963; Patrick & Reimer, 1966, 1975; Krammer & Lange-Bertalot, 1986, 1988, 1991a, b). There were forty-nine taxa known only from North America, fifteen known only from the New World, and 381 that could not be identified to the species level and possibly are undescribed taxa with limited distribution.

Only the 433 benthic diatom taxa that reached relative abundance of at least 1% in at least six samples were included in the data set of 2735 samples. Of these taxa, 390 (90%) are known from several continents. There were seven taxa found only in North and South America, and seventeen found only in North America. The nineteen undescribed taxa are possibly limited to North America.

Table 2 lists the thirty nonplanktonic taxa with the highest number of occurrences in 2735 samples. All these taxa are cosmopolitan and very common in continental waters.

Similarity of assemblage composition among ecoregions

Results of the DCA of the 2735 sample data set showed great heterogeneity in diatom assemblage composition (Fig. 2). The length of the gradient of the first ordination

axis was 7.3 SD units. The second axis was 7.6 SD units long. These values indicate high species turnover and are typical for sets of samples from diverse habitats. The small differences between the first four eigenvalues (0.6, 0.6, 0.57, 0.4) suggests that no single gradient in species composition was much stronger than all others. The DCA diagram shows centroids of ecoregions (Fig. 2a) and species (Fig. 2b) in the ordination space of the first and second axes. A group of ecoregions on the right side of the diagram (Ecoregion 9 and all those further to the right) is characterized by relatively arid climate and low relief, except ecoregions 10.2 and 10.4 that have mountainous relief, but were represented by only four sites. Diatom species showing maximum abundance in samples from ecoregions clustered on the right side of the diagram are known to prefer silty sediments and high alkalinity water: examples include *Nitzschia reversa* W. Smith, *N. capitellata* Hustedt, *N. agnita* Hustedt, *Craticula accomoda* (Hustedt) D.G. Mann, *Pleurosigma salinarum* Grunow, *Tryblionella levidensis* W. Smith, and others (Fig. 2b). Northern and mountainous ecoregions clustered in the left lower quadrant of the diagram (Ecoregions 6, 5, 8.4, 5.2). Diatoms common in these ecoregions are mostly epilithic, preferring fast-flowing and cold waters. Several species of *Epithemia* and *Cymbella* are characteristic of those areas. Separation of ecoregions and species along the first axis indicate that this axis is most probably

Taxon name	% of samples where taxon was encountered
<i>Achnantheidium minutissimum</i> (Kützing) Czarnecki	59.7
<i>Rhoicosphenia curvata</i> (Kützing) Grunow ex Rabenhorst	42.6
<i>Gomphonema parvulum</i> (Kützing) Kützing	41.1
<i>Navicula minima</i> Grunow	41.0
<i>Navicula cryptotenella</i> Lange-Bertalot	39.8
<i>Amphora pediculus</i> (Kützing) Grunow	39.4
<i>Cocconeis placentula</i> var. <i>euglypta</i> (Ehrenberg) Cleve	37.8
<i>Nitzschia inconspicua</i> Grunow	36.3
<i>Nitzschia amphibia</i> Grunow	33.2
<i>Nitzschia palea</i> (Kützing) W. Smith	32.2
<i>Encyonema minutum</i> (Hilse in Rabenhorst) D.G. Mann	32.0
<i>Reimeria sinuata</i> (Gregory) Kociolek and Stoermer	31.0
<i>Navicula gregaria</i> Donkin	29.4
<i>Gomphonema pumilum</i> (Grunow) Reichard and Lange-Bertalot	29.1
<i>Navicula capitatoradiata</i> Germain	27.5
<i>Cocconeis placentula</i> var. <i>lineata</i> (Ehrenberg) Van Heurck	26.8
<i>Achnanthes lanceolata</i> (Breb. in Kützing) Grunow	26.4
<i>Melosira varians</i> Agardh	24.3
<i>Navicula germainii</i> Wallace	24.0
<i>Nitzschia frustulum</i> (Kützing) Grunow	23.4
<i>Nitzschia dissipata</i> (Kützing) Grunow	22.6
<i>Navicula tripunctata</i> (O.F. Muller) Bory	22.5
<i>Staurosirella pinnata</i> (Ehrenberg) Williams and Round	22.5
<i>Achnanthes rostrata</i> Østrup	22.1
<i>Fragilaria vaucheriae</i> (Kützing) Peterson	20.8
<i>Nitzschia archibaldii</i> Lange-Bertalot	19.8
<i>Cocconeis pediculus</i> Ehrenberg	18.1
<i>Nitzschia palea</i> var. <i>debilis</i> (Kützing) Grunow	17.0
<i>Cymbella affinis</i> Kützing	16.8
<i>Synedra ulna</i> (Nitzsche) Ehrenberg	16.7

Table 2 Diatom taxa with highest occurrence in National Water Quality Assessment (NAWQA) data set

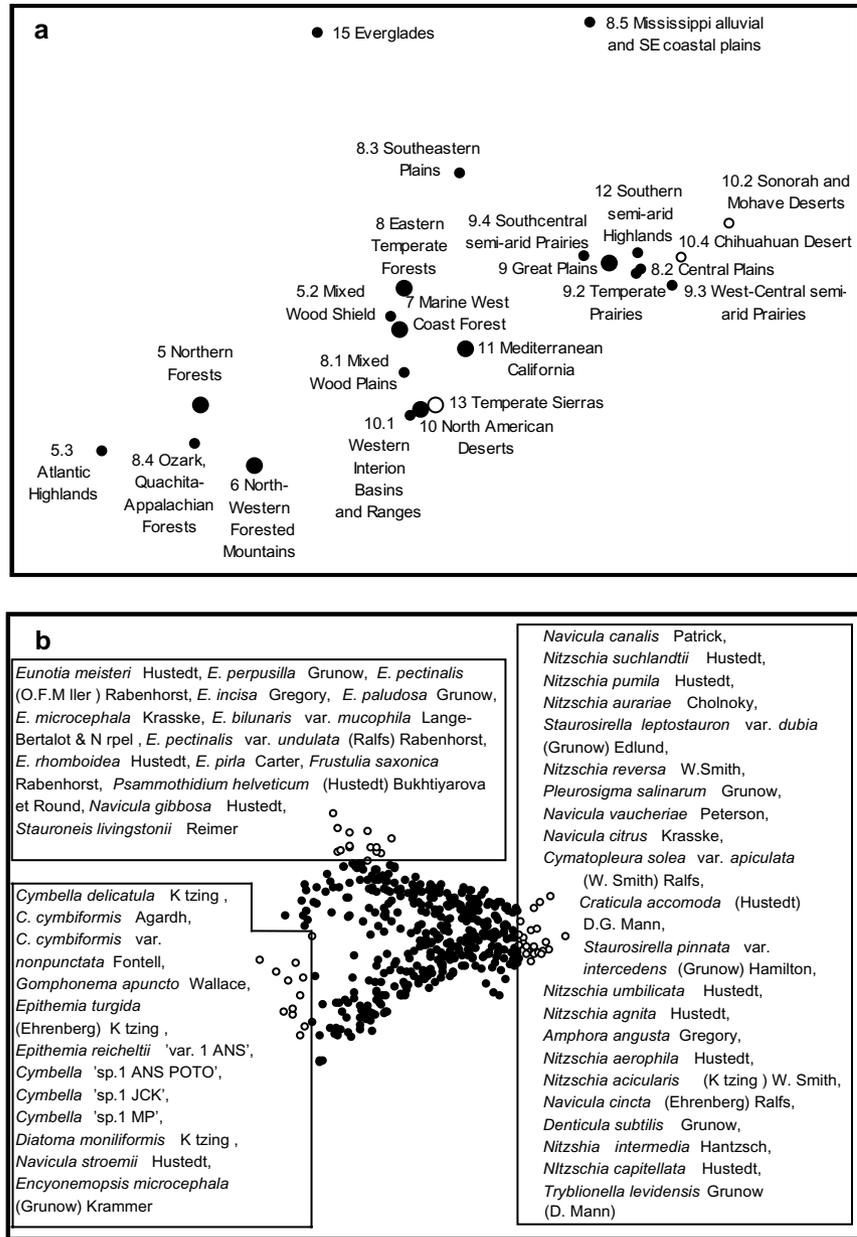


Figure 2 Detrended correspondence analysis of 2735 sample data set: plots showing position of ecoregions (a) and diatom taxa (b) in the ordination plane spanned by the first (horizontal) and the second (vertical) DCA axes. (a) Big circles: Level 1 Ecoregions, small circles; Level 2 Ecoregions, filled circles; ecoregions with more than five sites, empty circles; ecoregions with five sites or less. (b) taxa names in the boxes correspond to the centroids contained within the same boxes and shown by empty circles.

associated with an underlying gradient of slope/current velocity and water temperature. The first axis does not, however, represent a geographical gradient as adjacent ecoregions did not cluster together along it. Ecoregions in the upper part of the diagram (Florida Everglades and adjacent coastal and Mississippi plains) have a warm and wet climate; they have lowland rivers and canals that drain vast marshy areas. Corresponding communities are rich in *Eunotia* and other acidophilic diatoms. The second axis is most probably related to a pH gradient, but as the ecoregions located in the south-eastern part of the USA were clustered together along this axis, it could also represent a geographical pattern.

Relative importance of geographical vs. environmental factors in explaining diatom assemblage structure

We compared the relative influence of spatial and environmental factors on the structure of diatom communities by means of partial CCA (Table 3). The total variation explained by the model (TVE) was very low in the two large data sets comprised of all ecoregions. It was less than 12% in the 582 sample data set, and increased to 19% in the 247 sample data set that included only one sample per site. The TVE greatly increased for data sets of smaller geographical extent (up to 84% in the data set representing the 'Northern Forests' ecoregion). Because the numbers of

samples and taxa in those data sets were lower, the increase in the TVE was most probably the result of two factors that often contribute to higher TVE: narrowing of geographical area and decreased noise (Økland, 1999).

The fraction of the TVE attributed to spatial variables increased with geographical area. In Level 2 Ecoregions it was 15–22% of TVE, compared with 23–31% of TVE in Level 1 Ecoregions and the all-ecoregions data sets. The environmental variables explained a considerably higher proportion of variance in species data than spatial variables, especially in Level 2 Ecoregions.

Major environmental gradients influencing species distributions

The relative importance of environmental parameters in explaining species composition of diatom assemblages varied among geographical areas. Ordination diagrams in Fig. 3 show results of the CCAs with environmental variables as constraints. Graphs on the left side represent the results of simple CCAs; those on the right show the results of partial CCAs with spatial covariables. If an environmental variable appears to be important in the simple CCA, but not in the partial CCA, it is spatially structured. This variable may still have an effect on the diatom assemblage structure, but this effect is hard to distinguish from the influence of geographical factors. We compared results of the analyses of four data sets: one data set of 582 RTH samples from all ecoregions and the three regional data sets that had the highest number of samples: Ecoregions 6, 8, and 9.

In partial CCAs of the 582 RTH sample data set, twenty environmental variables were used as constraints, and all nine spatial variables as covariables. The fraction of the variation in the species data explained by environmental parameters was 8.5%, and only 7.3% when spatial variation was taken into account. The first ordination axis in both

analyses was related to a complex gradient that has oligotrophic fast-flowing rivers at one end and eutrophic lowland rivers at the other. Substrate type was strongly related to this axis in simple CCA (Fig. 3a). This is because samples in lowland rivers were often taken from wood substrates because rocks were difficult to find. The second axis in both cases expressed a gradient in pH and mineral content. The effect of air and water temperature seems to be strong in simple CCA (Fig. 3a), but not in partial CCA (Fig. 3b). This is logical because temperature variation is strongly spatially structured. The arrow representing the variable 'mean annual air temperature' (MAT) is much longer in the diagram showing results of the simple CCA (Fig. 3a) than in that of the partial CCA (Fig. 3b). This is because MAT is strongly spatially structured, and its influence is difficult to distinguish from 'pure' latitudinal variation.

The major gradients in the 'East Temperate Forests' (8) data set are similar to those in the all-ecoregions data set. This is probably because most samples in the all-ecoregion data set are from this large ecoregion. In the simple CCA, the first axis corresponds to a gradient of smaller, mountainous rivers to larger, lowland rivers (Fig. 3c). When spatial variation is partialled out, the nutrient gradient also corresponds to the first axis (Fig. 3d). The second ordination axis, as in the case of the all-ecoregions data set, represents a clear alkalinity/pH gradient. In Fig. 3c, arrows MAT and pH are both long and point in almost opposite directions. This is because most rivers in the south-eastern coastal plain have a low pH. However, when spatial variables are factored out (Fig. 3d), the effect of MAT becomes negligible, but that of pH is expressed strongly. Unlike air temperature, pH acts at the finer spatial scale and is often controlled by local features.

Ordination axes for the 'Great Plains' (9) ecoregion are more difficult to interpret, probably because of the uneven

Table 3 Variation in species data explained by two blocks of variables in partial canonical correspondence analysis (CCA)

Data set	No. of samples	No. of taxa in the data set	Total explained variation (TVE) (% of total variance)	Spatial parameters (% of TVE)	Environmental parameters (% of TVE)	Interaction of spatial and environmental parameters (% of TVE)
All ecoregions	582	377	11.9	28.3	57.7	14.0
All ecoregions	247	248	18.9	29.0	57.9	13.1
5 Northern Forests	26	64	83.7	30.9	43.6	25.5
6 North-western Forested Mountains	38	59	74.6	29.7	51.4	18.9
8 Eastern Temperate Forests	306	215	21.1	25.9	56.7	17.4
9 Great Plains	99	151	41.1	24.2	54.6	21.2
10 NA Deserts	68	74	50.4	23.4	67.2	9.4
8.1 Mixed Wood Plains	44	120	56.8	16.7	72.7	10.6
8.2 Central USA Plains	49	113	62.7	15.4	66.7	17.9
8.3 South-eastern USA Plains	86	174	36.8	21.9	67.5	10.6
8.4 Ozark, Quachita-Appalachian Forests	96	156	39.1	16.6	69.6	13.8
8.5 Mississippi alluvial and south-eastern coastal plain	31	121	72.8	21.8	61.9	16.3

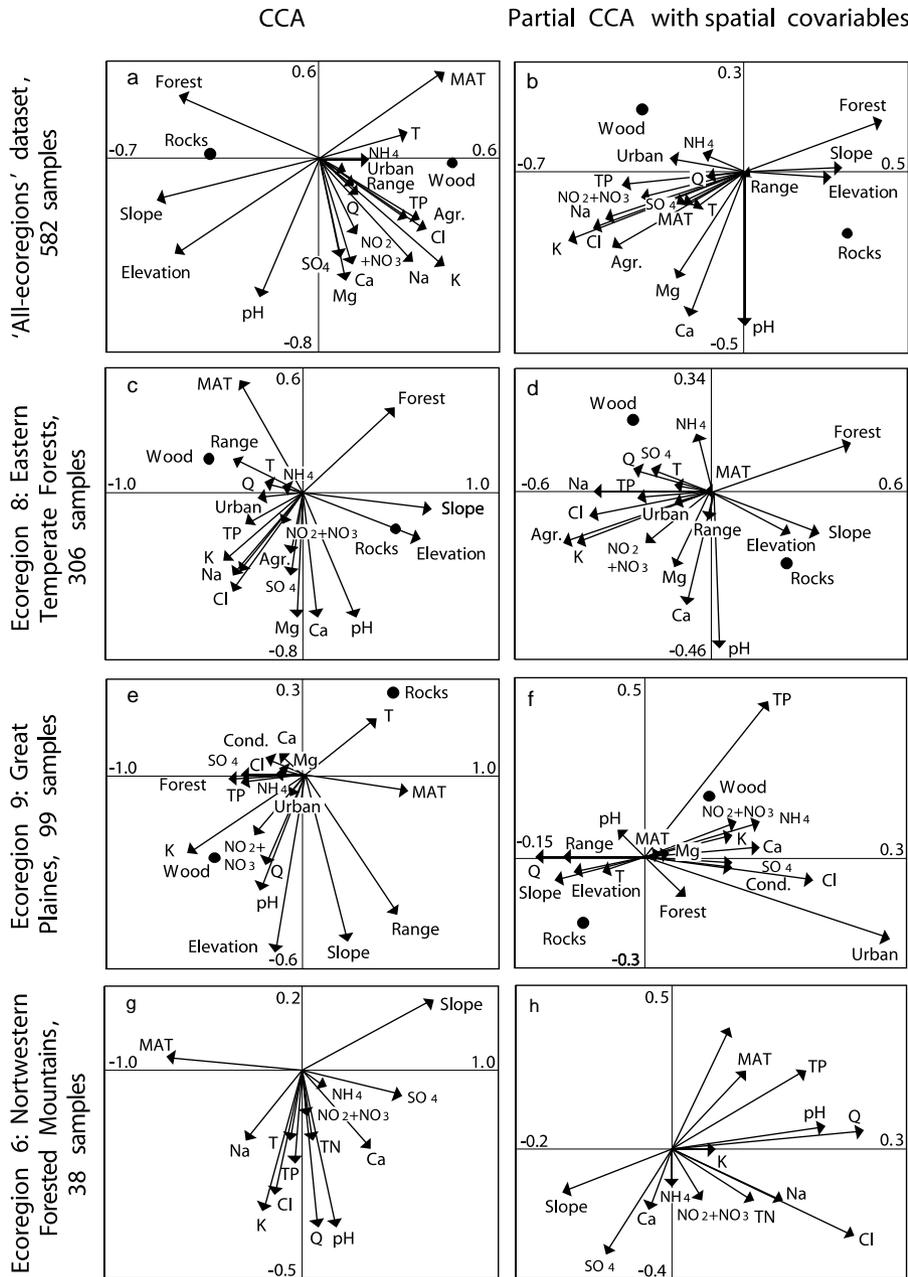


Figure 3 Results of canonical correspondence analysis without covariables (left column) and with spatial parameters as covariables (right column): plots of environmental variables in the ordination space of the first (horizontal) and the second (vertical) ordination axes. Arrows correspond to scores of continuous variables and filled circles are centroids of nominal variables.

distribution of sampling sites, which are clumped into several aggregates across this ecoregion. The first axis of the simple CCA (Fig. 3e) corresponds to latitude: mean annual air temperature is obviously higher in the South. Total pressure and $[\text{NH}_4^-]$ are also showing a 'northern' trend, most probably not reflecting any regular pattern, but because some of the northern sites happened to have high concentrations of these nutrients. The second axis corres-

ponds to a mountains-to-plains gradient. When spatial variables are factored out (Fig. 3f), air temperature also loses its importance. The strongest variation in diatom composition is now connected to a complex gradient of mountains/low mineral content/low nutrients-plains/higher mineral content/high nutrients.

The first CCA axis is closely associated with mean annual air temperature in the Ecoregion 'North-western Forested

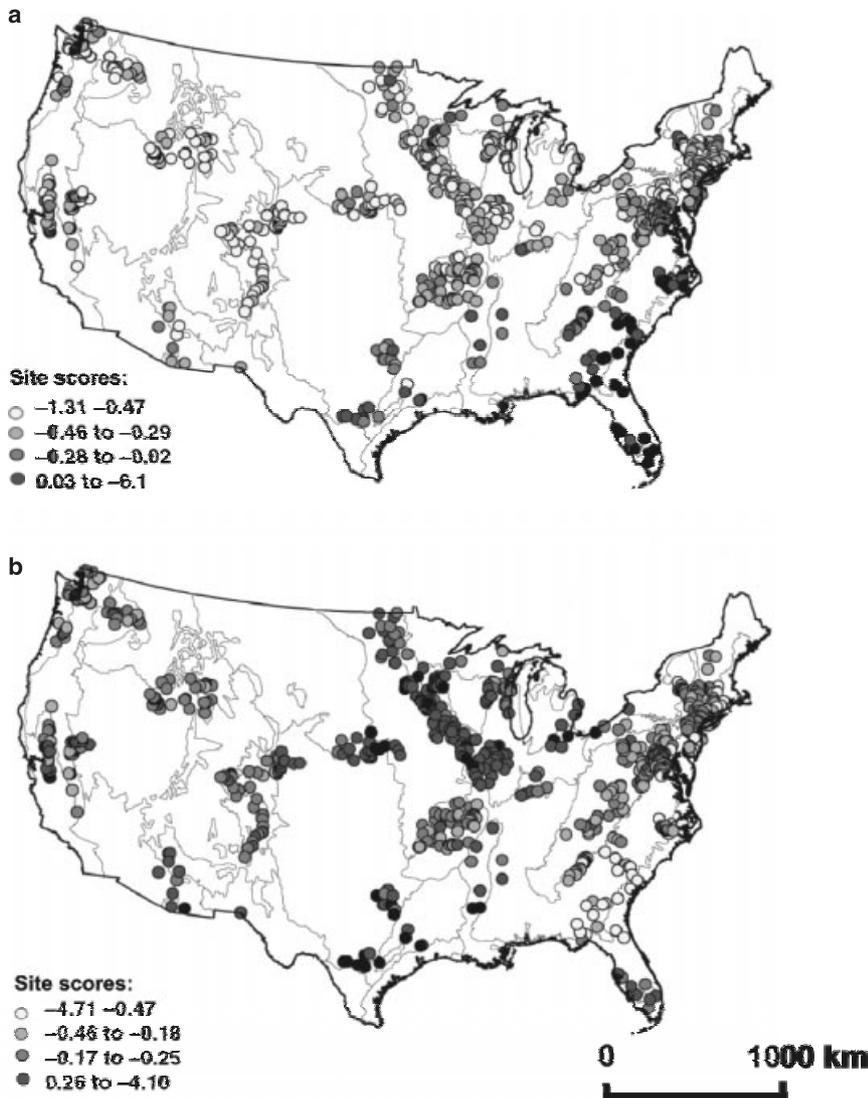


Figure 4 Site scores from CCA of the 2735 sample data set with nine spatial parameters as constraining variables along the first (a) and the second (b) ordination axes. Light shading indicate low, and dark shading – high site scores.

Mountains' (6) ordination. This reflects differences in diatom communities between southern and northern parts of this ecoregion (Fig. 3g). The second axis corresponds to the gradient of large/more alkaline/eutrophic and small/less buffered/oligotrophic rivers. Taking away spatially structured variability does not completely remove the effect of temperature (Fig. 3h), because in mountainous areas temperature also strongly depends on the elevation.

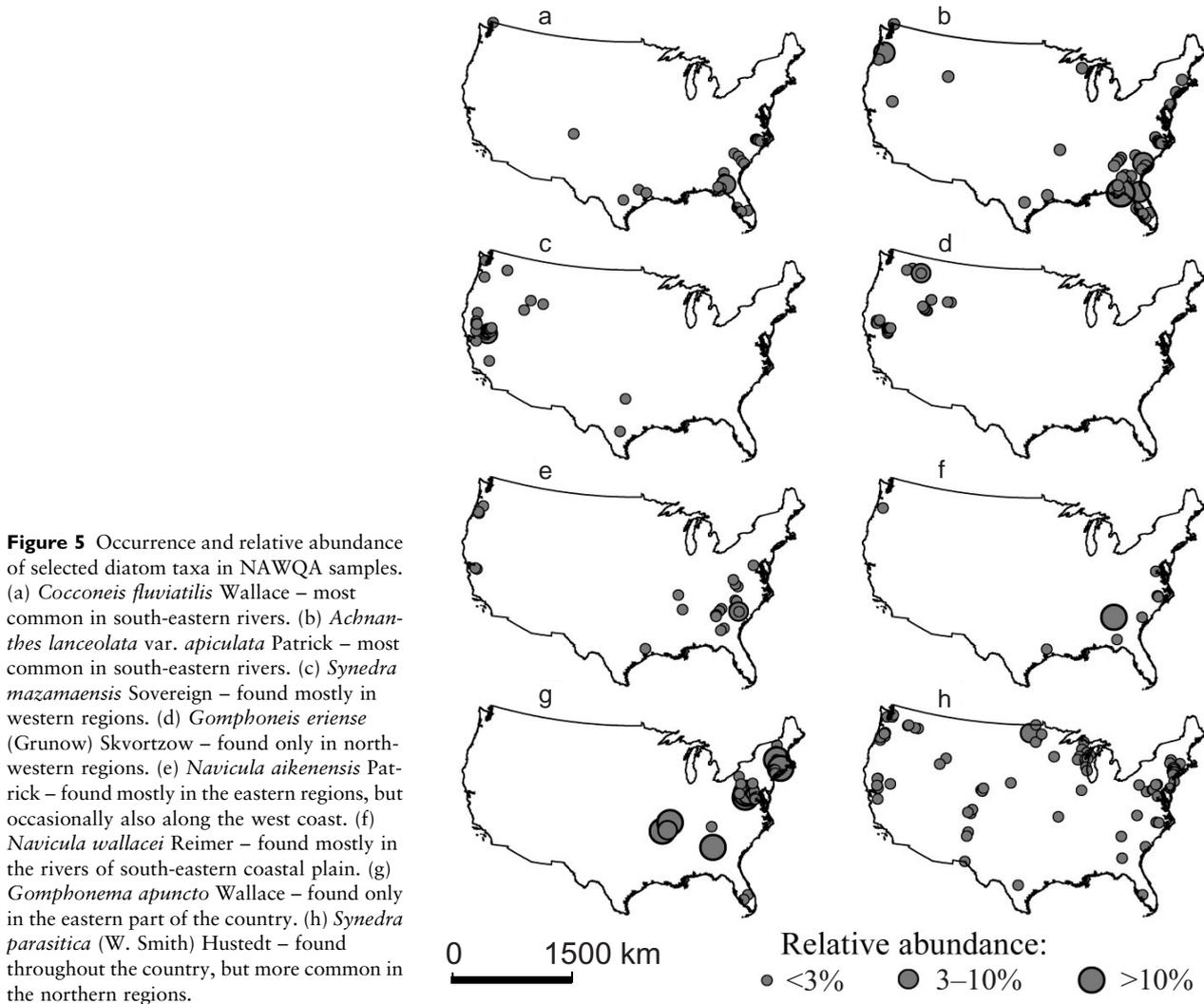
Geographic patterns of species distributions

The variance partitioning results (Table 3) show that environmental parameters consistently explain more variation than spatial factors. Still, at the national scale, the fraction of the spatially structured variation reaches almost 30% of TVE, and thus shows that the distribution of at least some species must be limited by geographical/dispersal factors. To reveal the main geographical patterns in species composition

we ran CCA with nine spatial parameters as constraining variables on the 2735 sample data set. Figure 4 is a map that shows first CCA axis (A) and second axis (B) site scores. The darker shades indicate higher, and lighter shades indicate lower ordination scores for the corresponding sites. The sites along the south-eastern coast received the highest scores along the first ordination axis, while most of the sites with lowest scores were located in interior Western Mountains [Ecoregions 'Western Cordillera' (6.2) and 'Western Interior Basin and Ranges' (10.1)] (Fig. 4a). Species that scored high on the first CCA axes were *Fragilaria synegrotasca* Lange-Bertalot, *Synedra delicatissima* W. Smith, *Nitzschia amphibia* var. *frauenfeldii* Grunow, *N. nana* Grunow, *A. exigua* var. *elliptica* Hustedt, *Gomphonema gracile* Ehrenberg, *A. grana* Hohn and Hellerman, *Terpsinoe musica* Ehrenberg, *Cocconeis fluviatilis* Wallace, *A. lanceolata* var. *apiculata* Patrick. Occurrence and relative abundance in NAWQA samples of the last two species are shown in Fig. 5a,b.

Species with the lowest scores on the first CCA axis included *C. cymbiformis* var. *nonpunctata* Fontell, *C. laevis* Naegeli ex. Kützing, *N. sinuata* var. *delognei* (Grunow) Lange-Bertalot, *N. solita* Hustedt, *A. dau*i Foged, undescribed species of *Nitzschia* (*Nitzschia* 'cf. *tubicola* Grunow') and *Navicula* (*Navicula* 'sp. 7 ANS WCR', *N.*'sp.8 ANS WCR'), *Hannaea arcus* (Ehrenberg) Patrick, *S. mazamaensis* Sovereign, and *Gomphoneis eriense* (Grunow) Skvortzow. Occurrence and relative abundance of the last two species is shown in Fig. 5c,d. The second CCA axis clearly separates sites of the eastern US coast (lower scores) and Texas (high scores). Sites of the intensive agricultural areas in the Upper Midwest [Ecoregions 'Temperate Prairies' (9.2) and 'Central USA Plains'(8.2)] also received high scores along the second CCA axis. Species with the highest scores along the second axis, and consequently more abundant on the eastern coast, were mostly acidophilic diatoms such as *Neidium densestriatum* (Østrup) Krammer, *Neidium alpinum* Hustedt, *Eunotia naegeli*i Migula, *E. microcephala* Krasske ex. Hustedt,

E. rhomboidea Hustedt, *E. pectinalis* var. *undulata* (Ralfs) Rabenhorst, *E. meisteri* Hustedt, *E. pirla* Carter and Flower, *E. monodon* Ehrenberg, *N. aikenensis* Patrick, *N. wallacei* Reimer. Occurrence and relative abundance of the last two species are shown in Fig. 5e,f. Species that received low scores on the second axis are characteristic of waters of high mineral content: *Denticula subtilis* Grunow, *Amphora angusta* Gregory, *Pleurosigma salinarum* Grunow, *Nitzschia vitrea* Norman, *Mastogloia smithii* Thwaites, *Diploneis pseudovalis* Hustedt, *D. smithii* var. *pumila* (Grunow) Hustedt, *Achnanthes exilis* Kützing, *Terpsinoe musica* Ehrenberg, *Gyrosigma nodiferum* (Grunow) Reimer, *Tryblionella levidensis* W. Smith, *N. texana* Patrick, *N. sanctae-crucis* Østrup, *Fallacia omissa* (Hustedt) D.G. Mann, and some undescribed species of *Achnanthisdium* and *Encyonopsis* found only in Texas. The result of this analysis demonstrates that the major gradients of species composition do not simply go in north–south or east–west directions. In fact, the geographical patterns are far more complex. As we did



not factor out variation explained by environmental parameters, the pattern revealed by this analysis is partly influenced by the environment. To find species whose distribution is better explained by spatial than environmental variables we used the 'all-ecoregions' data set of 582 RTH samples to run partial CCA. The nine spatial parameters were used as constraining variables. The covariables were the nineteen land-use, chemical and physical parameters; mean annual air temperature was not included because of its high correlation with latitude. The species placed by CCA at the extremes of the first four ordination axes were considered as having distribution more closely following some spatial pattern than being determined strictly by environmental factors. *Cocconeis fluviatilis* (Fig. 5a), *N. aikenensis* (Fig. 5e), and *N. wallacei* Reimer (Fig. 5f), were common mostly in the south-east. *Synedra mazamaensis* (Fig. 5c) occurred mostly in the western part of USA. *Navicula paucivittata* Patrick and *G. rhombicum* Fricke exhibited a more complex distribution pattern, characterized by rarity of occurrence in the central plains. Many other species are revealed by this analysis as having spatial distributions restricted to small geographical areas. These include *N. amphibia* var. *frauenfeldii* and *Fragilaria synegetesca* that occurred mostly in South Florida and adjacent areas, undescribed species of *Epithemia* from the Ozark Plateau, *Achnanthisidium*, *Navicula*, *Encyonopsis* and *Gomphonema* from Texas, *Navicula* from Nebraska and Colorado, and several species of *Gomphonema* from different localities in the western interior.

None of the analyses revealed a clear east–west floristic gradient. Some species, however, were more or less widely distributed only in the western part of US, as for example *Gomphoneis eriense* (Fig. 5d), or in the eastern part of the country, as *G. apuncto* Wallace (Fig. 5g).

Using the geographical coordinates for all 2735 quantitative samples, we identified the taxa most common at sampling sites in the northern and southern parts of the USA. We calculated weighted means of the taxa along the latitudinal gradient and then ranked taxa accordingly to these 'latitude optima.' The taxa that were more common in the North and that occurred in at least ten samples were: *G. minutiforme* Lange-Bertalot and Reichardt, *S. parasitica* W. Smith (Fig. 6h), *S. parasitica* var. *subconstricta* (Grunow) Hustedt, *S. famelica* Kützing, *Gomphoneis eriense* (Fig. 6d), *Encyonema prostratum* (Berkeley) Kützing, *Diatoma mesodon* (Ehrenberg) Kützing, *N. tantula* Hustedt, *G. rhombicum*, and two not yet described species of *Cymbella*. The examples of taxa with 'southern' affinity and occurrence in at least ten samples are: *S. delicatissima* W. Smith, *N. amphibia* var. *frauenfeldii* Grunow, *N. obtusa* W. Smith, *N. liebethuthii* Rabenhorst, *N. aerophila* Hustedt, *A. exigua* var. *elliptica* Hustedt, *A. exigua* var. *heterovalva* Krasske, *A. amoena* Hustedt, *Terpsinoe musica* Ehrenberg, *G. gracile* Ehrenberg, and *Amphora angusta* Gregory.

Species indicating nutrient enrichment

The relationships between diatom species and nutrients are influenced by complex ecological gradients. Figure 6 com-

pare species-environment biplots resulting from two CCAs of the 'all-ecoregions' data set of 582 samples. In the first CCA we used TN and TP as the only constraining variables (Fig. 6a). In the second CCA we also used TN and TP as constraints, but in addition we entered pH, slope and all nine spatial parameters as covariables (Fig. 6b). We chose these covariables because the results of previous CCAs showed their importance as factors underlying the structure of diatom communities. The TN and TP arrows point in the same direction along the first axis in both analyses. This indicates that the first ordination axis corresponds to a general nutrient enrichment gradient. In the partial CCA, there is greater dispersion of species centroids (Fig. 6b). This shows that in partial CCA, separate effects of TP and TN were more distinguishable; diatoms responded to varying concentrations and ratios of N and P and not just to a generalized nutrient gradient.

The species whose centroids are found at the extremes of the first ordination axes of each CCA are potential indicators of high or low concentrations of nutrients. We show approximate positions of selected species that were found in at least six samples and were placed by ordination at the extremes of the first CCA axis. There is considerable, but not complete, overlap between the results of the two analyses (Figs 6a,b). Comparison of lists of seventeen species placed at the left end of the first axis of both CCAs shows that only thirteen of them are common to both lists. Only ten species are common to two lists of species placed at the right end of the first CCA axis. These differences between the results of the two CCAs show that some species could be mistaken for indicators of high or low nutrients, while in fact, their distribution is controlled primarily by other factors that correlate with nutrient concentrations.

DISCUSSION

'Downstream' gradient

Both DCA and CCA show that the gradient of fast-flowing rivers of mountainous areas to lowland rivers is the most important in structuring diatom assemblages at the national scale. This is a complex gradient, because variables such as slope, elevation, concentration of nutrients, land-use and temperature are correlated. This complex gradient largely corresponds to the 'downstream' coenocline widely known because of the popularity of the 'River Continuum Concept' (Vannote *et al.*, 1980). The striking difference between diatom assemblages of headwaters and downstream reaches of rivers was recognized well before multivariate analysis was used to identify major environmental gradients in stream communities. Margalef (1960) characterized algal communities of headwaters as 'young', compared with the 'mature' communities' of the downstream reaches. This point of view reflects the general trend of discharge to be more stable as rivers get larger and slope decreases. Communities of headwaters are generally 'younger' because the frequency and magnitude of disturbance are higher. Nutrient enrichment is another general 'downstream' trend. The downstream

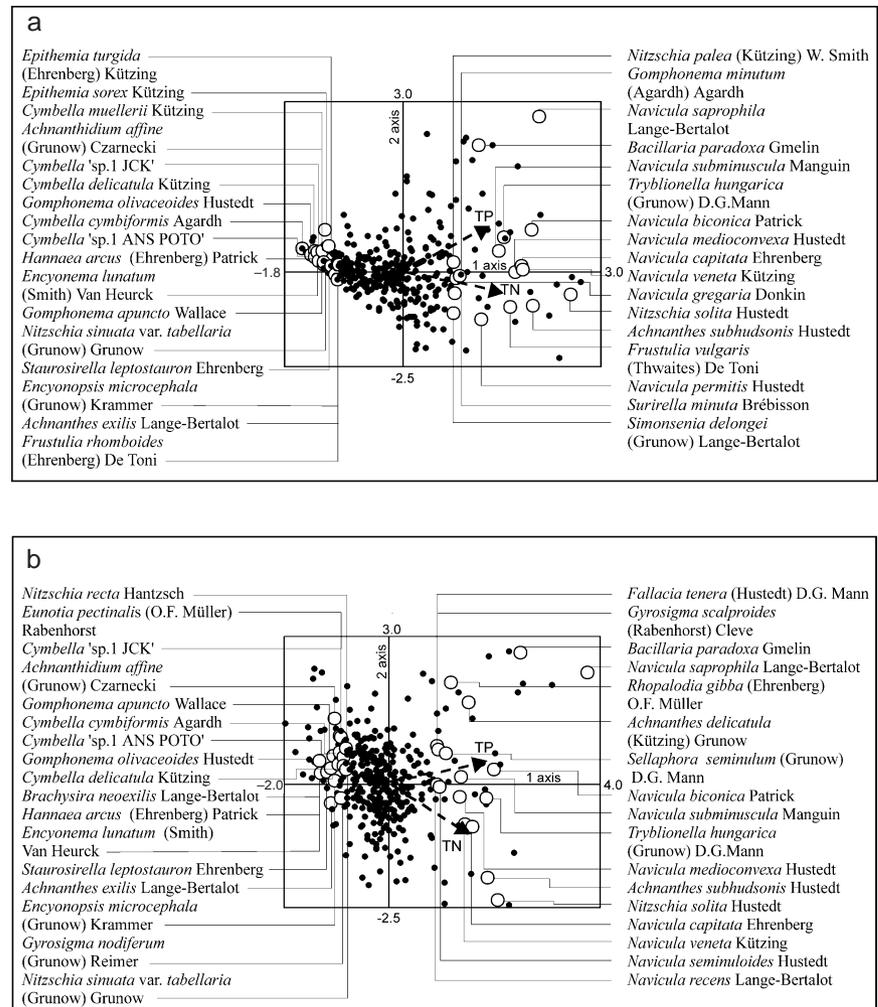


Figure 6 Biplots of environmental variables and diatom taxa in the ordination space of the first (horizontal) and the second (vertical) ordination axes from CCA of 582 RTH samples data set without covariables (a) and with pH, conductivity, slope and nine spatial parameters as covariables (b). Position of arrow heads correspond to scores of constraining environmental variables (TN and TP) and circles represent taxa scores.

enrichment with dissolved salts, and nutrients in particular, was especially clear in our analyses of the data sets representing large geographical areas. The first axis of DCA (Fig. 2) separated mountainous areas from plains, and, respectively, epilithic oligotrophic diatoms from predominantly epipelagic alkaliphilic and eutrophic species. Canonical correspondence analysis (Fig. 3a,b) shows directly that this 'mountains-to-plains' gradient and associated gradients of substrate, land-use and eutrophication are of great importance for diatom assemblages at the national scale.

Effect of ionic strength and pH

The overriding effect of water chemistry, especially salinity and pH, on species composition of diatom assemblages has been recognized since the beginning of ecological studies of this algal group. These relationships were first documented in the 'Halobion system' (Kolbe, 1927, 1932) and the pH-classification system of Hustedt (1938–39). Many studies that employed multivariate analyses to analyse river algal communities showed that they varied conspicuously along

gradients of ionic concentration and pH (Johansson, 1982; Leclercq & Depiereux, 1987; Pan *et al.*, 1996).

At the broad national scale, some regions of the US have rivers with much higher mineral content than other regions. Climate, particularly precipitation, is a major cause of these differences. Our analyses of the diatom data show an obvious 'eastern coast–western interior' gradient related to climate. The humid areas of the eastern coast of the United States, especially the south-eastern coastal plains, support a diatom flora rich in acidophilic taxa. The arid areas of the western part of the country have abundant alkaliphilic diatoms (Fig. 2). At smaller spatial scales, the factors governing water mineral content other than climate become more evident. These include bedrock type, land-use and watershed size. As pH is partly regulated by the water buffering capacity or amount of alkaline metals, it also depends on all the above-mentioned factors, but also on the extent of wetlands which are related to local topographical features. The relative importance of a parameter in explaining variation among diatom assemblages also strongly depends on its range of variation in the data set. The pH

range in the 'Great Plains' (9) data set was 7.1–8.8, compared with 5.6–9.0 in the 'Eastern Temperate Forests' (8) data set. Accordingly, the influence of pH was more pronounced in the latter data set (compare length of the pH arrow at Fig 3d,f). pH is consistently found to be one of the most influential parameters in areas with relatively low water buffering capacity (Carpenter & Waite, 2000), but not in areas where all sites are high in water mineral content, which is the case of the upper Illinois (Leland & Porter, 2000). In data sets with a relatively large range of pH, as for example in the Appalachian Mountains (Pan *et al.*, 1996), it is hard to build effective phosphorus-inference models because diatom communities respond primarily to pH variation. On the other hand, polluted rivers are often rich not only in nutrients, but in other ions too. This was the case for data sets analysed by Leland (1995) and Leland & Porter (2000). The response of diatom communities to nutrient enrichment seemed to be evident, but in fact the CCA shows that the effect of nutrients and conductivity on species composition was practically indistinguishable. In these circumstances care should be taken not to mistake the effect of salinity for the effect of nutrient enrichment.

Effect of temperature

At the national scale, diatom assemblages clearly responded to variation in temperature. Notably, MAT explained more variation in diatom assemblages than water temperature (Fig. 3). This finding was surprising, because diatom species are known to respond to seasonal changes in water temperature (Moore, 1977a, b), and we would expect a better response to the water temperature measured at the time of algal sampling than to such a crude estimation as average annual air temperature. We suppose that one-time measurements of water temperature provided a poor assessment of temperature regime in rivers because of large diurnal and day-to-day oscillations, and this is perhaps one of the reasons why the relationship between air temperature and diatom assemblages appeared to be stronger in CCAs. We could expect a stronger effect of the water temperature if it was estimated in a more accurate way, for example, if it was monitored constantly during a period preceding algal sampling.

DeNicola (1996) reviewed autoecological and field studies of temperature effects on freshwater periphyton. He concluded that the 'relative importance of temperature on geographical distribution is unknown', despite the relatively abundant information that exists on the temperature requirements of many diatom species and patterns of temporal and spatial distribution of species within individual rivers or watersheds. Although Pienitz *et al.* (1995), Weckström *et al.* (1997), and Rosén *et al.* (2000) successfully used diatom assemblages to infer water temperature in lakes, temperature is rarely found as an important environmental variable in regional data sets. This is because it is rare for temperature to vary in a region more than any other factor, or it correlates with other environmental factors and their effects are difficult to separate (Anderson, 2000). At the broad geographical scale, a pronounced effect of tempera-

ture can be expected because of climatic variation. The contiguous US lies almost entirely in the temperate climatic zone, with the exception of tropical southern Florida, but the range of MAT in this study was as wide as 1.5–23.4 °C (factor of 15.6). A good response of diatom assemblages to the variation of MAT gives us the hope that large-scale studies such as the NAWQA programme will provide important information on temperature requirements of diatom species.

Spatial patterns

Large-scale spatial patterns of species dispersal, independent of local environmental characteristics, are rarely taken into account in ecological studies of diatom communities. Presumably, this influence is considered negligible when the geographical area studied is small. However, spatial dispersal patterns cannot be neglected when national-scale data are analysed. Our analyses show that almost one-third of the explainable variation in diatom species composition at the national scale can be attributed to spatial factors. Even if some part of this 'spatial variation' is the result of unmeasured environmental parameters, such a high number indicates that distributions of at least some taxa were better explained by spatial rather than environmental parameters. In terms of diatom species composition, the south-eastern coastal plains, including Florida, stand out from the rest of the US. However, none of the diatom taxa that played an important role in the river benthic communities were found to be limited to this area. All taxa that showed affinity to south-eastern regions were also encountered elsewhere. *Cocconeis fluviatilis* occurred mostly in the rivers of south-eastern coastal plains, but occasionally was found in other areas (Fig. 6a). It was also recorded in the Great Lakes (Stoermer & Kreis, 1978), Utah (Rushforth & Squires, 1985; Rushforth & Merkley, 1989), Oklahoma (Kennedy & Pfister, 1984), and Ohio (Kline & Lowe, 1975). A similar pattern of distribution is found for *N. aikenensis*, *N. wallacei*, and *A. lanceolata* var. *apiculata*. The tendency of these taxa to occur predominantly in the South-east cannot be explained by measured environmental parameters, in contrast to the distribution of several species of *Eunotia* and *Neidium* which are acidophilic taxa and were also most common in south-eastern rivers that often have low pH.

Some species, such as *G. eriense* (Fig. 6d), and some undescribed *Navicula* and *Nitzschia*, were found in this study only in western regions of the US. *Synedra mazamaensis* was recorded mostly in the west, but occasionally also in Texas. Kocielek & Stoermer (1988) showed that *G. eriense* was once more widely distributed than at present, and documented its disappearance from the Great Lakes as a result of eutrophication. This is an example of restricted distribution resulting from habitat loss.

A south-to-north gradient in species composition could be explained by two mechanisms: temperature requirements of diatom species and dispersal rates that limit the pool of species available for colonization. *Gomphonema gracile* showed a 'southern' affinity in our data set, but has been

reported as being cosmopolitan, and was described by Ehrenberg from Germany. Krammer & Lange-Bertalot (1986) note that this species is abundant in the tropics and Northern Europe. In a data set of lakes from the Canadian Arctic (Pienitz *et al.*, 1995), *G. gracile* had one of the lowest temperature optima. The real cause of this surprising distribution remains to be studied. Occurrence of *A. amoena* predominantly in the south corresponds well to the known abundance of this species in the tropics (Hustedt, 1952; Krammer & Lange-Bertalot, 1991b; Chang, 1992). *Achnanthes exigua* var. *heterovalva* and *A. exigua* var. *elliptica* have been characterized as cosmopolitan taxa (Hustedt, 1938–39, 1959), but showed ‘southern’ trend in this study.

Some of the taxa that showed ‘northern’ affinity in NAWQA data set, such as *Diatoma mesodon*, are well known to prefer cold waters. *Synedra famelica* is possibly a ‘nordic-alpine’ species according to Hustedt (1959). Other taxa that showed affinity toward the northern areas, such as *S. parasitica*, *S. parasitica* var. *subconstricta*, *G. minutiforme*, *G. rhombicum*, *G. eriense*, *Encyonema prostratum*, and *N. tantula*, may also have relatively low temperature optima or their distribution can be explained by other factors that correlate with latitude.

As we expected, the spatial variables explained a higher proportion of community variation in data sets representing large geographical areas than smaller areas (Table 3). The amount of the total variation explained by models (TVE) was very low in data sets with many samples from large geographical areas. Low TVE is often a feature of large data sets, and does not necessarily mean that important environmental factors have been overlooked. Even though the TVE is low, the proportion that is explained by different sets of environmental parameters is informative (Økland, 1999).

Our analysis shows that at any spatial scale, there were species with patchy distributions, not easily explained by environmental factors. Some proportion of this ‘spatial’ variation can be related to unmeasured environmental parameters that were spatially structured. Another part of it is due to spatial autocorrelation, or the tendency of geographically close sites to have similar flora (Jongman *et al.*, 1995).

Mann & Droop (1996) note that ‘detection of biogeographical patterns and assessment of rarity are inextricably linked with taxonomy’. That is, different patterns of distribution can be detected depending on the taxonomy used in a study. If a wider species concept is applied, species with restricted distribution tend to be overlooked. A finer-scale taxonomy leads to recognition of more species with narrow geographical and ecological distributions. For the above reason Kociolek & Spaulding (2000) argue against the view that most diatoms are cosmopolitan. Indeed, lumping of species with similar morphologies into one category might give the wrong impression about species distributions. However, even the most sophisticated methods of morphological analysis only raise hypotheses; they cannot provide proof that entities with slightly different morphology are isolated reproductively and are separate species, not morphotypes of the same species. In this study we sometimes had to apply a rather broad species concept because of limitations of having several

people involved in the counting. Nevertheless, approximately one quarter of the taxa found in 2735 NAWQA samples are known to, or might possibly have, distributions limited to the Americas. However, the proportion of these taxa decreases to 4% when only the relatively abundant and frequent river diatoms are considered. The list of the most frequently occurring species (Table 2) consists of only taxa with cosmopolitan distributions. This might be because the common species are those that are adapted to the most common habitat, which, in the case of the United States and Europe, is a nutrient-enriched moderately alkaline river of the temperate climatic zone. Abundance of this habitat in one continent means higher chances of dispersal for its inhabitants to other parts of the world. Vice versa, rarity of certain habitats, for example oligotrophic sites, lowers dispersal chances for the organisms living in and adapted to those conditions. Another possibility is that cosmopolitan morphospecies consist of more than one biological species. In the absence of strong scientific evidence on the reproductive barriers between morphologically close groups of diatoms, practicing ecologists must rely on morphological evidence alone and are limited to working with entities that may not be separate biological species, but rather aggregations of related species. Future taxonomic work based on the combination of molecular and morphological approaches (Manhart & McCourt, 1992) will further elucidate the relative importance of environmental and dispersal/colonization constraints in structuring diatom assemblages.

Implications of broad-scale distribution patterns on the use of diatoms as indicators of eutrophication

Large and taxonomically consistent data sets such as the one analysed in this study can potentially provide a basis for development of diatom-based metrics for water-quality assessment, and in particular assessment of eutrophication. Modern use of diatoms as indicators of environmental conditions is based on the development of transfer functions using a variety of numerical procedures. One of the requirements for the successful development and application of those models is that ‘environmental variables, other than the one of the interest, have negligible influence, or their joint distribution with the variable of interest in the fossil set is the same as in the training set’ (Birks *et al.*, 1990). This requirement cannot always be satisfied, and techniques such as WA-PLS were introduced to circumvent the presence of strong secondary gradients in the training data set (Birks, 1998). Our analyses show that presence of gradients collinear with, or orthogonal to, the nutrient gradient can influence placement of species along the trophic scale. Strong orthogonal gradients, for example the pH-gradient in the data sets with large pH range, obscure the effect of nutrients. In the case of collinear gradients, the effect of each of them is difficult to distinguish. Development of reliable diatom metrics will require careful investigation of data sets in terms of the major environmental gradients underlying species composition, and additional sampling aimed to increase community variation along the gradient of interest.

Differences in species composition among geographical areas, caused not only by environmental variation, but also by historical processes of species dispersal and colonization, add difficulties in applying uniform methods for water quality assessment. Diatoms that are identified as the same species using the broad species concept currently employed in most popular floras, could, in fact, be separate biological species, having different physiological requirements. The chances are higher that these 'cryptic' species inhabit different geographical areas rather than the same area because the higher frequency of allopatric compared with sympatric speciation. Taking into account these considerations, we recommend that diatom-based metrics should first be developed for limited geographical areas with the most uniform environmental characteristics possible.

CONCLUSIONS

Although the references used in this study for taxonomic identifications are based on a generally wide species concept, 26% of all recorded taxa were not described in the literature. Most of these taxa were rare, but the great proportion of undescribed taxa suggest the substantial need for future taxonomic work.

Multivariate analyses of NAWQA data sets revealed that the following ecological gradients explained most of the variation in benthic diatom assemblages at the national scale: (1) the gradient of fast-flowing and mostly oligotrophic rivers of mountain areas to lowland, mostly eutrophic rivers, (2) water mineral content and pH and (3) temperature, which also correlates strongly with latitude and altitude. In the data set that included sites from the whole country, measured environmental parameters explained roughly twice as much variation in species data as compared with spatial variables. The relative contribution of spatial vs. environmental factors decreased with geographical extent of the data sets. Effects of individual factors that constitute complex ecological gradients can be difficult to discern. Presence of environmental factors correlated with nutrient concentrations should be taken into account when species–nutrient relationships are studied, because these internal correlations of environmental factors in a data set can lead to incorrect conclusions about species nutrient requirements. The considerable proportion of variation in diatom assemblages explained by spatial parameters alone suggests that some historical processes of species dispersal play an important role in shaping diatom assemblages. Considering the complexity of factors affecting species composition of benthic diatoms, it is advisable to develop diatom-based metrics for nutrient estimation for limited geographical areas, preferably ecologically uniform in terms of factors other than nutrients.

ACKNOWLEDGMENTS

This research was supported by the US Geological Survey NAWQA programme through a cooperative agreement with

the Academy of Natural Sciences. Carol Couch and David Muller helped to obtain environmental data. Loren Bahls, Todd Clason, William Cody, Kalina Manoylov, Lont Marr, Eduardo Morales, Charles Reimer and Noma-Ann Roberts analysed the NAWQA diatom samples. John Kingston significantly improved taxonomic quality of the counts. Loren Bahls, Todd Clason, William Cody, Rex Lowe, Kalina Manoylov, Lont Marr, Eduardo Morales, Charles Reimer and Jan Stevenson contributed to our efforts to combine diatom counts in a single taxonomically consistent data set. Pat Cotter and Frank Acker participated in the creation of the computer applications and data base that made possible management of the data used in this study. Carol Couch, David Hart, John Smol, and an anonymous reviewer provided useful comments on the manuscript.

REFERENCES

- Anderson, N.J. (2000) Diatoms, temperature and climatic change. *European Journal of Phycology*, **35**, 307–314.
- Bahls, L.L. (1982) Eight new diatom genus records from Montana. *Proceedings of the Montana Academy of Science*, **41**, 79–86.
- Bahls, L.L. (1983) A new diatom from southeastern Montana. *Proceedings of the Montana Academy of Science*, **42**, 1–6.
- Bahls, L.L., Weber, E.E. & Jarvie, J.O. (1984) Ecology and distribution of major diatom ecotypes in the southern Fort Union coal region of Montana. *Geological Survey Professional Paper*, **1289**, 1–151.
- Benson, C.E. & Rushforth, S.R. (1975) The algal flora of Huntington Canyon, Utah, U.S.A. *Bibliotheca Phycologica*, **18**, 1–177.
- Birks, H.J.B. (1998) Numerical tools in palaeolimnology – Progress, potentialities, and problems. *Journal of Paleolimnology*, **20**, 307–332.
- Birks, H.J.B., Line, J.M., Juggins, S., Stevenson, A.C. & ter Braak, C.J.F. (1990) Diatoms and pH reconstruction. *Palaeolimnology and lake acidification* (eds R.W. Battarbee, J. Mason, I. Renberg and J. F. Talling), pp. 37–52. Proceedings of a Royal Society Discussion Meeting Held on 25 August 1989. The Royal Society, London.
- Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.
- ter Braak, C.J.F. (1996) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology*, **67**, 1167–1179.
- ter Braak, C.J.F. & Šmilauer, P. (1998) *CANOCO reference manual and user's guide to CANOCO for Windows: software for canonical community ordination* Version 4. Microcomputer Power, Ithaca.
- Camburn, K.E., Kingston, J.C. & Charles, D.F. (1984–86) *PIRLA diatom iconograph*. PIRLA Unpublished Report Series, Report 3. Department of Biology, Indiana University, Bloomington.
- Camburn, K.E., Lowe, R.L. & Stoneburner, D.L. (1978) The haptobenthic diatom flora of Long Branch Creek, South Carolina. *Nova Hedwigia*, **30**, 149–279.

- Carpenter, K.D. & Waite, I.R. (2000) Relations of habitat-specific algal assemblages to land-use and water chemistry in the Willamette basin, Oregon. *Environmental Monitoring and Assessment*, **64**, 247–257.
- Chang, T.-P. (1992) On *Achnanthes amoena* Hustedt. *Diatom Research*, **7**, 397–402.
- Clark, R.L. & Rushforth, S.R. (1977) Diatom studies of the headwaters of Henrys Fork of the Snake River, Island Park, Idaho, U.S.A. *Bibliotheca Phycologica*, **33**, 1–204.
- Czarnecki, D.B. & Blinn, D.W. (1978) Diatoms of the Colorado River in Grand Canyon National Park and Vicinity (Diatoms of Southwestern USA. II). *Bibliotheca Diatomologica*, **38**, 1–181.
- DeNicola, D.M. (1996) Periphyton responses to temperature at different ecological levels. *Algal ecology: freshwater benthic ecosystems* (eds R. J. Stevenson, M. I. Bothwell and R. L. Lowe), pp. 149–181. Academic Press, Inc, San Diego.
- Edlund, M.B. (1994) Additions and confirmations to the algal flora of Itasca State Park. II. Diatoms from Chambers Creek. *Journal of the Minnesota Academy of Science*, **59**, 10–21.
- Eloranta, P. & Kwandrans, J. (1996) Testing the use of diatoms and macroalgae for river quality monitoring in Finland. *Use of algae for monitoring rivers II* (eds A.B. Whitton and E. Rott), pp. 119–124. Institut für Botanik, Universität Innsbruck, Innsbruck.
- Gurtz, M.E. (1993) Design of biological components of the National Water-Quality Assessment (NAWQA) Program. *Biology monitoring of aquatic systems* (eds S. L. Loeb and A. Spacie), pp. 323–354. Lewis Publishers, Boca Raton.
- Hansmann, E.W. (1973) Diatoms of the streams of Eastern Connecticut. State Geological and Natural History Survey of Connecticut. *Department of Environmental Protection*, **106**, 1–119.
- Hohn, M.H. & Helleman, H. (1963) The taxonomy and structure of diatom populations from three eastern North American rivers, using three sampling methods. *Transactions of the American Microscopical Society*, **82**, 250–329.
- Hustedt, F. (1930a) Bacillariophyta (Diatomaceae). *Die Süßwasser-Flora Mitteleuropas* (ed. A. Pascher), pp. 1–466. Heft 10, Part 1–2, G. Fisher, Jena.
- Hustedt, F. (1930b) Die Kieselalgen Deutschlands, Österreichs und der Schweiz. 1. *Kryptogamen-Flora Von Deutschland, Österreichs und der Schweiz*, Vol. 7 (ed. L. Rabenhorst), pp. 1–920. Akademische Verlagsgesellschaft, Leipzig.
- Hustedt, F. (1938–39) Systematische und ökologische Untersuchungen über die Diatomeen-Flora von Java, Bali, Sumatra nach dem Material der Deutschen Limnologischen Sunda-Expedition. III. Die ökologischen Faktoren und ihr Einfluß auf die Diatomeenflora. *Archiv für Hydrobiologie*, **16** (Suppl.), 274–394.
- Hustedt, F. (1952) Neue und wenig bekannte Diatomeen IV. *Botaniska Notiser*, **4**, 366–410.
- Hustedt, F. (1959) Die Kieselalgen Deutschlands, Österreichs und der Schweiz. 2. *Kryptogamen-Flora Von Deutschland, Österreichs und der Schweiz*, Vol. 7 (ed. L. Rabenhorst), pp. 1–845. Akademische Verlagsgesellschaft, Leipzig.
- Hustedt, F. (1961–66) Die Kieselalgen Deutschlands, Österreichs und der Schweiz. 3. *Kryptogamen-Flora Von Deutschland, Österreichs und der Schweiz*, Vol. 7. (ed. L. Rabenhorst), pp. 1–816. Akademische Verlagsgesellschaft, Leipzig.
- Jackson, D.C. & Lowe, R.L. (1978) Valve ultrastructure of species of the diatom genera *Gyrosigma* and *Pleurosigma* from the Portage River drainage system, Ohio. *Transactions of the American Microscopical Society*, **97**, 569–581.
- Johansson, C. (1982) Attached algal vegetation in running waters of Jämtland, Sweden. *Acta Phytogeographica Suecica*, **71**, 1–84.
- Jongman, R.H.G., ter Braak, C.J.F. & van Tongeren, O.F.R. (1995) *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge.
- Juggins, S. & ter Braak, C.J.F. (1992) *CALIBRATE – a program for species-environment calibration by [weighted-averaging] partial least squares regression*. Environmental Change Research Centre, University College, London.
- Keithan, E.D. & Lowe, R.L. (1985) Primary productivity and spatial structure of phytolith growth in streams in the Great Smoky Mountains National Park, Tennessee. *Hydrobiologia*, **123**, 59–67.
- Kennedy, M.T. & Pfister, L. (1984) New additions to the diatom (Bacillariophyceae) flora of Oklahoma. *Proceedings of the Oklahoma Academy of Science*, **64**, 37–39.
- Kline, P.A. & Lowe, R.L. (1975) Phytoplankton of the Sandusky River (Fremont Area). *Proceedings of the Sandusky river basin symposium* (eds D. B. Baker, W. B. Jackson and B. L. Prater), pp. 175–208. Bowling Green State University, Bowling Green, OH.
- Kociolek, J.P. & Kingston, J.C. (1999) Taxonomy, ultrastructure, and distribution of some gomphonemoid diatoms (*Bacillariophyceae: Gomphonemataceae*) from rivers in the United States. *Canadian Journal of Botany*, **77**, 686–705.
- Kociolek, J.P. & Spaulding, S.A. (2000) Freshwater diatom biogeography. *Nova Hedwigia*, **71**, 223–241.
- Kociolek, J.P. & Stoermer, E.F. (1988) Taxonomy, ultrastructure and distribution of *Gomphoneis herculeana*, *G. eriense* and closely related species (Naviculales: Gomphonemataceae). *Proceedings of the Academy of Natural Sciences of Philadelphia*, **140**, 24–97.
- Kociolek, J.P. & Stoermer, E.F. (1990) A new, highly variable *Gomphonema* (Bacillariophyceae) species from the Laurentian Great Lakes. *Ouvrage dédié à la mémoire du Professeur H. Germain* (eds M. Ricard and M. Coste), pp. 139–144. Koeltz, Koenigstein.
- Kociolek, J.P. & Stoermer, E.F. (1991) Taxonomy and ultrastructure of some *Gomphonema* and *Gomphoneis* taxa from the upper Laurentian Great Lakes. *Canadian Journal of Botany*, **69**, 1557–1576.
- Kociolek, J.P., Stoermer, E.F. & Edlund, M.A. (1995) Two new freshwater diatom species. *A century of diatom research in North America: a tribute to the distinguished careers of Charles W. Reimer and Ruth Patrick* (eds J. P. Kociolek and M. J. Sullivan), pp. 9–19. Koeltz USA, Champaign.
- Kolbe, R.W. (1927) Zur Ökologie, Morphologie und Systematik der Brackwasser-Diatomeen. *Pflanzenforschung*, **7**, 1–146.
- Kolbe, R.W. (1932) Grundlinien einer allgemeinen Ökologie der Diatomeen. *Ergebnisse der Biologie*, **8**, 221–348.
- Krammer, K. & Lange-Bertalot, H. (1986) Bacillariophyceae. 1. Teil: Naviculaceae. *Süßwasserflora von Mitteleuropa 2/1* (eds J. Ettl, J. Gerloff, H. Heynig and D. Mollenhauer), pp. 1–876. G. Fisher, Stuttgart.

- Krammer, K. & Lange-Bertalot, H. (1988) Bacillariophyceae. 2. Teil: Bacillariaceae, Epithemiaceae, Surirellaceae. *Süßwasserflora von Mitteleuropa 2/2* (eds J. Ettl, J. Gerloff, H. Heynig and D. Mollenhauer), pp. 1–596. G. Fisher, Stuttgart.
- Krammer, K. & Lange-Bertalot, H. (1991a) Bacillariophyceae. 3. Teil: Centrales, Fragilariaceae, Eunotiaceae. *Süßwasserflora von Mitteleuropa 2/3* (eds J. Ettl, J. Gerloff, H. Heynig and D. Mollenhauer), pp. 1–576. G. Fisher, Stuttgart.
- Krammer, K. & Lange-Bertalot, H. (1991b) Bacillariophyceae. 4. Teil: Achnantheaceae. *Süßwasserflora von Mitteleuropa 2/4* (eds J. Ettl, J. Gerloff, H. Heynig and D. Mollenhauer), pp. 1–437. G. Fisher, Stuttgart.
- Krejci, M.E. & Lowe, R.L. (1987) Spatial and temporal variation of epipsammic diatoms in a spring-fed brook. *Journal of Phycology*, **23**, 585–590.
- Kristiansen, J. (1996) Dispersal of freshwater algae – a review. *Hydrobiologia*, **336**, 151–157.
- Lange-Bertalot, H. & Moser, G. (1994) *Brachysira*. Monographie der Gattung. Wichtige Indikator-Species für das Gewässer-Monitoring und *Naviculadicta* nov. General Ein Lösungsvorschlag zu dem Problem *Navicula* sensu lato ohne *Navicula* sensu stricto. *Bibliotheca Diatomologica*, **29**, 1–212.
- Lawson, L.L. & Rushforth, S.R. (1975) The diatom flora of the Provo River, Utah, U.S.A. *Bibliotheca Phycologica*, **17**, 1–149.
- Leclercq, L. & Depiereux, E. (1987) Typologie des rivières oligotrophes du massif Ardennais (Belgique) par l'analyse multivariée de relevés de diatomées benthiques. *Hydrobiologia*, **153**, 175–192.
- Legendre, P. & Legendre, L. (1998) *Numerical ecology*. Elsevier, New York.
- Leland, H.V. (1995) Distribution of phytobenthos in the Yakima River basin, Washington, in relation to geology, land use, and other environmental factors. *Canadian Journal of Fisheries and Aquatic Sciences*, **52**, 1108–1129.
- Leland, H.V. & Porter, S.D. (2000) Distribution of benthic algae in the upper Illinois River basin in relation to geology and land use. *Freshwater Biology*, **44**, 279–301.
- Lowe, R.L. (1972) Notes on Iowa diatoms X: new and rare diatoms. *Proceedings of the Iowa Academy of Science*, **79**, 66–69.
- Lowe, R.L. & Kociolek, J.P. (1984) New and rare diatoms from Great Smoky Mountains National Park. *Nova Hedwigia*, **39**, 465–476.
- Manhart, J.R. & McCourt, R.M. (1992) Molecular data and species concept in the algae. *Journal of Phycology*, **28**, 730–737.
- Mann, D.G. & Droop, S.J.M. (1996) Biodiversity, biogeography and conservation of diatoms. *Hydrobiologia*, **336**, 19–32.
- Margalef, R. (1960) Ideas for a synthetic approach to the ecology of running waters. *Internationale Revue der Gesamten Hydrobiologie*, **45**, 133–153.
- Millie, D.F. & Lowe, R.L. (1981) Diatoms new to Ohio and the Laurentian Great Lakes. *Ohio Journal of Science*, **81**, 195–206.
- Moore, J.W. (1977a) Seasonal succession of algae in a eutrophic stream in southern England. *Hydrobiologia*, **53**, 181–192.
- Moore, J.W. (1977b) Seasonal succession of algae in rivers. II. Examples from Highland Water, a small woodland stream. *Hydrobiologia*, **80**, 160–171.
- Økland, R.H. (1999) On the variation explained by ordination and constrained ordination axes. *Journal of Vegetation Science*, **10**, 131–136.
- Omernik, J.M. (1995) Ecoregions: a spatial framework for environmental management. *Biology assessment and criteria: tools for water resource planning and decision making* (eds W. S. Davis and T. P. Simon), pp. 49–62. Lewis, Boca Raton, Florida.
- Pan, Y., Stevenson, R.J., Hill, B.H. & Herlihy, A.T. (2000) Ecoregions and benthic diatom assemblages in Mid-Atlantic Highlands streams, USA. *Journal of the North American Benthological Society*, **19**, 518–540.
- Pan, Y., Stevenson, R.J., Hill, B.H., Herlihy, A.T. & Collins, G.D. (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., Stevenson, R.J., Hill, B.H., Kaufmann, P.R. & Herlihy, A.T. (1999) Spatial patterns and ecological determinants of benthic algal assemblages in Mid-Atlantic streams, USA. *Journal of Phycology*, **35**, 460–468.
- Patrick, R. (1945) A taxonomic and ecological study of some diatoms from the Pocono Plateau and adjacent regions. *Farlowia*, **2**, 143–221.
- Patrick, R. (1946) Diatoms from Parschke Bog, Texas. *Notulae Naturae of the Academy of Natural Sciences of Philadelphia*, **170**, 1–7.
- Patrick, R. (1958) Some nomenclature problems and a new species and a new variety in the genus *Eunotia* (Bacillariophyceae). *Notulae Naturae of the Academy of Natural Sciences of Philadelphia*, **312**, 1–15.
- Patrick, R. (1959) New species and nomenclatural changes in the genus *Navicula* (Bacillariophyceae). *Proceedings of the Academy of Natural Sciences of Philadelphia*, **111**, 91–108.
- Patrick, R. (1996) *Rivers of the United States. V. III. The Eastern and Southeastern United States*. John Wiley & Sons, Inc, New York.
- Patrick, R. & Reimer, C. (1966) The diatoms of the United States. Vol. 1. *Monographs of The Academy of Natural Sciences of Philadelphia*, no. 13.
- Patrick, R. & Reimer, C. (1975) The diatoms of the United States. Vol. 2. Part 1. *Monographs of The Academy of Natural Sciences of Philadelphia*, no. 13.
- Patrick, R. & Roberts, N.-A. (1979) Diatom communities in the Middle Atlantic states. *Nova Hedwigia, Beiheft*, **64**, 265–283.
- Pienitz, R., Smol, J.P. & Birks, H.J.B. (1995) Assessment of freshwater diatoms as quantitative indicators of past climatic change in the Yukon and Northwest Territories, Canada. *Journal of Paleolimnology*, **13**, 21–49.
- Porter, S.D., Cuffney, T.F., Gurtz, M.E. & Meador, M.R. (1993) *Methods for collecting algal samples as part of the National Water Quality Assessment Program*. Open-File Report 93–409, US Geological Survey, Raleigh, NC [http://water.usgs.gov/nawqa/protocols/OFR-93-409/algp1/html].
- Reichardt, E. (1997) Taxonomische Revision des Artenkomplexes um *Gomphonema pumilum* (Bacillariophyceae). *Nova Hedwigia*, **65**, 99–129.
- Reimer, C.W. (1966) Consideration of fifteen diatom taxa (Bacillariophyta) from the Savannah River, including seven

- described as new. *Notulae Naturae of Academy of Natural Sciences of Philadelphia*, 397, 1–15.
- Rosén, P., Hall, R., Korsman, T. & Renberg, I. (2000) Diatom transfer-functions for quantifying past air temperature, pH and total organic carbon concentration from lakes in northern Sweden. *Journal of Paleolimnology*, 24, 109–123.
- Round, F.E. (1981) *The ecology of algae*. Cambridge University Press, New York.
- Rushforth, S.R. & Merkley, G.S. (1988) Comprehensive list by habitat of the algae of Utah, USA. *Great Basin Naturalist*, 48, 154–179.
- Rushforth, S.R. & Squires, L.E. (1985) New records and comprehensive list of the algal taxa of Utah Lake, Utah, USA. *Great Basin Naturalist*, 45, 237–254.
- Simonsen, R. (1987) *Atlas and catalogue of the diatom types of Friedrich Hustedt*. J. Cramer, Berlin.
- Stevenson, R.J. (1983) Effects of current and conditions simulating autogenically changing microhabitats on benthic algal immigration. *Ecology*, 64, 1514–1524.
- Stevenson, R.J. (1990) Benthic algal community dynamics in a stream during and after a spate. *Journal of the North American Benthological Society*, 9, 277–288.
- Stevenson, R.J. (1997) Scale-dependent determinants and consequences of benthic algal heterogeneity. *Journal of the North American Benthological Society*, 16, 248–262.
- Stevenson, R.J., Bothwell, M.L. & Lowe, R.L. (eds) (1996) *Algal ecology: freshwater benthic ecosystems*. Academic Press, Inc., San Diego.
- Stevenson, R.J. & Hashim, S. (1989) Variation in diatom community structure among habitats in sandy streams. *Journal of Phycology*, 25, 678–686.
- Stevenson, R.J. & Peterson, C.G. (1989) Changes in benthic diatom (Bacillariophyceae) immigration with habitat characteristics and cell morphology. *Journal of Phycology*, 25, 120–129.
- Stevenson, R.J. & Peterson, C.G. (1991) Emigration and immigration can be important determinants of benthic diatom assemblages in streams. *Freshwater Biology*, 26, 295–306.
- Stevenson, R.J. & Pryfogle, P. (1975) A comparison of diatom flora of Sandusky River and Tymochtee Creek. *Proceedings of the Sandusky river basin symposium* (eds D. B. Baker, W. B. Jackson and B.L. Prater), pp. 209–231.
- Stoermer, E.F. & Kreis, R.G. (1978) Preliminary checklist of diatoms (Bacillariophyta) from the Laurentian Great Lakes. *Journal of the Great Lakes Research*, 4, 149–169.
- Stoermer, E.F. & Smol, J.P. (eds) (1999) *The diatoms: applications for the environmental and earth sciences*. Cambridge University Press, Cambridge.
- Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. & Cushing, C.E. (1980) The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 130–137.
- Wallace, J.H. (1960) New and variable diatoms. *Notulae Naturae of the Academy of Natural Sciences of Philadelphia*, 331, 1–8.
- Weckström, J., Korhola, A. & Blom, T. (1997) The relationship between diatoms and water temperature in thirty subarctic Fennoscandian lakes. *Arctic and Alpine Research*, 29, 75–92.

BIOSKETCHES

Marina Potapova is a phycologist at the Patrick Center for Environmental Research, Academy of Natural Sciences, Philadelphia. She has interests in ecology and taxonomy of algae, especially river diatoms.

Donald Charles is the leader of the Phycology Section, Patrick Center for Environmental Research, Academy of Natural Sciences, Philadelphia. His interests include aquatic ecology, paleoecology, and use of diatoms as environmental indicators, particularly to assess river water quality.