

# Response of algal metrics to nutrients and physical factors and identification of nutrient thresholds in agricultural streams

Robert W. Black · Patrick W. Moran ·  
Jill D. Frankforter

Received: 12 October 2009 / Accepted: 26 May 2010 / Published online: 25 June 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Many streams within the United States are impaired due to nutrient enrichment, particularly in agricultural settings. The present study examines the response of benthic algal communities in agricultural and minimally disturbed sites from across the western United States to a suite of environmental factors, including nutrients, collected at multiple scales. The first objective was to identify the relative importance of nutrients, habitat and watershed features, and macroinvertebrate trophic structure to explain algal metrics derived from deposition and erosion habitats. The second objective was to determine if thresholds in total nitrogen (TN) and total phosphorus (TP) related to algal metrics could be identified and how these thresholds varied across metrics and habitats. Nutrient concentrations within the agricultural areas were elevated and greater than published threshold values. All algal metrics examined responded to nutrients as hypothesized. Although nutrients

typically were the most important variables in explaining the variation in each of the algal metrics, environmental factors operating at multiple scales also were important. Calculated thresholds for TN or TP based on the algal metrics generated from samples collected from erosion and deposition habitats were not significantly different. Little variability in threshold values for each metric for TN and TP was observed. The consistency of the threshold values measured across multiple metrics and habitats suggest that the thresholds identified in this study are ecologically relevant. Additional work to characterize the relationship between algal metrics, physical and chemical features, and nuisance algal growth would be of benefit to the development of nutrient thresholds and criteria.

**Keywords** Nutrients · Agriculture · Algal metrics · Streams · Thresholds · Habitat

## Introduction

Nutrient effects on aquatic communities have widely been recognized as an issue of ecological concern (Hynes 1970; Belore et al. 2002; Gurbuz and Kivrak 2002; Heinz Center 2002; US Environmental Protection Agency 2002). The National Water Quality Inventory: 1998 Report to Congress (US Environmental Protection Agency 2002) cites nutrients (nitrogen and phosphorus)

---

R. W. Black (✉) · P. W. Moran  
US Geological Survey, Washington Water  
Science Center, 934 Broadway, Suite 300,  
Tacoma, WA 98402 USA  
e-mail: rwblack@usgs.gov

J. D. Frankforter  
US Geological Survey, Montana Water  
Science Center, 3162 Bozeman Avenue,  
Helena, MT 59601, USA

as the third leading cause of water quality impairment in the rivers, lakes, and estuaries of the United States. Of the 23% of the total river and stream miles assessed in that report, 30% were impaired due to nutrient enrichment. Agriculture was the leading anthropogenic factor contributing to 59% of the reported water quality problems in impaired rivers and streams. Recently, the Wadeable Streams Assessment (US Environmental Protection Agency 2006) determined that nitrogen and phosphorus were the two most prominent stressors of wadeable streams in the United States.

Although controlling nutrient inputs often is an effective management tool in agricultural watersheds, stream ecosystems respond to a complex suite of interacting conditions of which nutrient enrichment is but one. To advance the scientific basis for establishing or refining numeric nutrient regulatory standards, it is important to understand how natural physical conditions, macroinvertebrate trophic structure and anthropogenic disturbances interact with nutrients to influence stream algae and algal indicators. Many factors can regulate primary producers in streams, i.e., nutrient availability, hydrodynamics, grazing, turbidity, riparian shading, and anthropogenic affects, such as addition of toxic compounds, introduced species, and watershed development (Dodds and Welch 2000; Newall and Walsh 2005).

Previous studies have focused on the variation in algal biomass and community structure with varying nutrient concentrations (Dodds et al. 1997; Chételat et al. 1999; Soinenen et al. 2004) and the identification of other non-nutrient factors that may influence these relationship (Dodds et al. 2002). Light, climate, geology, land use, substrate size, and vegetation all have been identified as factors that can control algal populations in addition to nutrient conditions and related water chemistry parameters. In some cases, algal biomass and community structure are related more to factors measured at larger scales such as land use (Leland and Porter 2000; Snyder et al. 2002; Taylor et al. 2004) or patterns in hydrology (Biggs 1996; Biggs and Smith 2002) than to nutrients measured at local and ephemeral scales. Although broad-scale factors have been related to algal biomass and community structure, the hierarchical

arrangement of these factors has received less attention (Stevenson 1997). It has been suggested (Poff 1997) that large-scale factors are causal agents that constrain the expression of small-scale factors such as nutrient concentrations at a particular stream site. The broad suite of environmental factors that influence algal biomass and species composition makes the development of nutrient-algal predictive models difficult. However, the development of such models (e.g., nutrient-algal regression models) can provide valuable information necessary for water quality management decisions and the development of nutrient criteria (Dodds et al. 2002).

The present study focuses on the response of benthic algal communities collected from agriculturally dominated to low-impact sites from across the western United States to a suite of environmental factors, including nutrients, collected at multiple scales. The first objective was to identify the relative importance of nutrient concentrations, specific watershed and reach environmental factors, and macroinvertebrate trophic structure in explaining algal metrics derived from two habitat types. As part of this objective, the study examined the importance of nitrogen and phosphorus levels in relationship to other watershed, instream, and macroinvertebrate trophic measures. The second objective was to examine if thresholds in total nitrogen (TN) and total phosphorus (TP) related to the measured algal metrics could be identified and how these thresholds varied across algal metrics and habitat types.

## Materials and methods

### Study sites

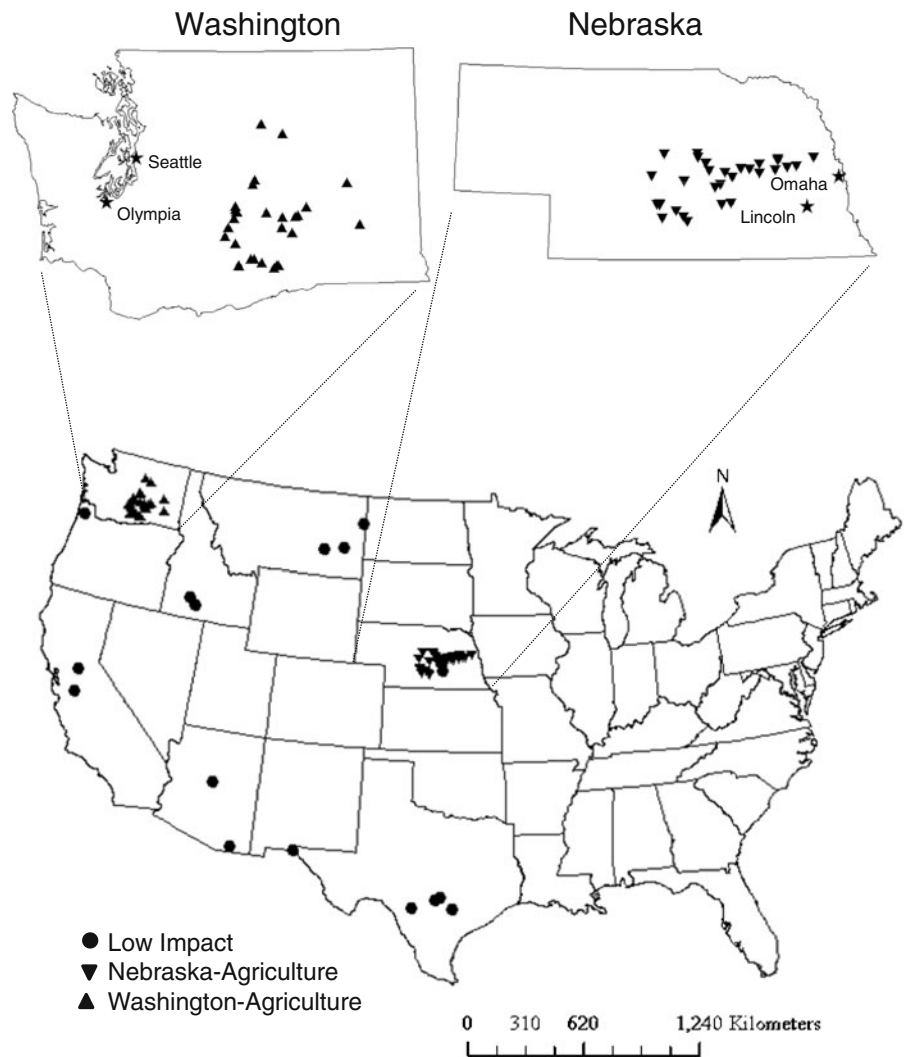
The study examined the water chemistry, algal and macroinvertebrate communities, and watershed and reach-scale habitat conditions from 73 sites throughout the western United States sampled as part of the US Geological Survey's (USGS) National Water Quality Assessment Program (NAWQA). Fifty-six sites were equally distributed between two agricultural regions in the arid to semiarid areas of the western United States where the natural hydrologic networks have been

largely altered to allow the growth of crops in areas with little natural precipitation. The agricultural study regions discussed below are the Central Columbia Plateau–Yakima River Basins in central Washington and the Central Nebraska Basins in central Nebraska (Fig. 1). To select sites with a gradient of potential nutrient conditions in Washington and Nebraska, nitrogen and phosphorus loading to all watersheds was modeled using a combination of fertilizer sales, land use, animal population, and atmospheric deposition (Ruddy et al. 2006). Based on existing nutrient data and modeled nutrient loading, sites were selected that spanned a range of nutrient conditions.

The Washington study area is predominantly arid with a climate that is subject to large seasonal and diurnal variations. The study area includes a large-scale, federally sponsored, irrigation system that irrigates more than one million acres of productive agricultural lands that receives 20–38 cm of rain per year. Detailed descriptions of the environmental setting of the Washington study area are available in Rinella et al. (1992), Jones and Wagner (1995), and Munn et al. (2002).

The Nebraska study area consists of the area drained by the Platte River Basin between the confluences of the North and South Platte Rivers near North Platte, Nebraska to its confluence with

**Fig. 1** Washington and Nebraska agricultural and low-impact sampling sites in the central and western United States



the Missouri River south of Omaha, Nebraska (Zelt and Jordan 1993). Annual precipitation in the study area ranges from about 46 cm in the west to 76 cm in the east (Frenzel et al. 1998). Land use is predominately rangeland in the west and cropland in the east, except where limited by topography (Zelt and Jordan 1993). Platte River stream flow is controlled by reservoir operations on the North Platte River in Wyoming and Lake McConaughy (since 1941) near the mouth of the North Platte River. The South Platte River in Nebraska is influenced primarily by irrigation return flows at low flows (Huntzinger and Ellis 1993).

The remaining 17 sites were located throughout the western United States and were characterized by very low levels (<10%) of agriculture and/or urbanization. These sites are part of the low-impact group (Fig. 1). The low-impact sites covered conditions detected across the western United States, but all the sites were in relatively dry climates similar to those in the agricultural study areas and were critical to establishing a gradient of potential nutrient conditions. Potential low-impact sites from the NAWQA database were excluded from this dataset if any mining activity or extensive anthropogenic activities were known of within the watershed. No consideration was given to nutrient loads or nutrient conditions prior to their selection as low-impact sites.

All the sites in this study were in the Central and Western Plains' diatom-based geographic region as determined by Potapova and Charles (2007). To create the Central and Western Plains' diatom region, Potapova and Charles (2007) performed nonmetric multidimensional scaling (NMS) on diatom data from 1,240 sites sampled throughout the United States as part of the NAWQA study. The goal of the NMS ordination was to delineate geographic regions that naturally have relatively homogeneous diatom assemblages and environmental settings (Potapova and Charles 2007).

#### Study reach habitat characterization

A study reach of approximately 20 to 30 wetted channel widths in length was established at each site. Reach-scale data were collected along five

to 11 transects distributed proportionally within the study reach. Habitat measures for each study reach included an assessment of instream and bank features such as water velocity, depth, width, substrate size, substrate embeddedness, channel bankfull width, bank vegetation cover, and estimates of canopy closure (made using a spherical densitometer). All width and depth measurements were to the nearest tenth of a meter. Instantaneous stream discharge was measured following standard USGS methods (Rantz 1982). Habitat characterizations were performed during low-flow conditions between 1993 and 2003. A detailed description of the reach-scale habitat and riparian delineation methods is available in Fitzpatrick et al. (1998) and Biggs and Kilroy (1994). Study reach habitat variables that were collected or derived and used in the data analysis are shown in Table 1.

#### Water quality/chemistry

Water chemistry samples were collected twice at all sites: once 30 days prior to the habitat characterization and once during the habitat assessments. Samples were collected using isokinetic depth-integrated equal-width increment sampling methods unless the stream was too shallow or water velocity was insufficient, in which case, samples were collected as multivertical grab samples (US Geological Survey 2006). Field water quality properties were measured during each sampling event and included water temperature, dissolved oxygen, specific conductance, and pH using laboratory meters that were calibrated daily prior to use. Water chemistry analysis included nutrients (TN and TP as well as dissolved species), alkalinity, and dissolved organic carbon. Samples also were collected for the determination of suspended sediment. Once a chemical sample was appropriately collected and processed, it was stored on ice and shipped to the laboratory within 24 h where they were maintained at 4°C prior to analysis (Patton and Gilroy 1998). All laboratory analyses for chemical constituents were conducted at the USGS National Water Quality Laboratory in Denver, CO, using methods by Fishman (1993) and Patton and Kryskalla (2003). Nutrient and dissolved organic carbon samples were analyzed

**Table 1** Predictor variables by scale and variable type

Predictor variables	Mean	(Min., Max.)
<b>Watershed characteristics</b>		
Mean annual basin runoff in 2002 (mm/year)	216.2	(1.5, 1163.1)
Base flow index in the basin (base flow divided by total flow, %; Wolock 2003)	54.9	(24.2, 81.1)
Min. soil permeability (in/h)	1.6	(0.3, 7.1)
Rainfall and runoff factor (“R factor” of universal soil loss equation)	91.7	(2.8, 286.7)
Percent clay	19.6	(4.3, 44.0)
Percent silt	45.3	(9.9, 72.8)
Percent sand	35.1	(4.1, 85.8)
Mean watershed elevation (m)	1052.8	(240.3, 3033.1)
Max. water capacity for soil layer horizon (in. of water/in. of soil depth)	0.2	(0.1, 0.2)
Land surface slope (%)	12.9	(1.3, 38.5)
Watershed area (km <sup>2</sup> )	758.7	(3.2, 6378.8)
Soil material < 0.074 mm (% by weight)	52.5	(14.1, 95.7)
Soil material < 2 mm (% by weight)	79.6	(31.2, 100)
Soil material < 5 mm (% by weight)	83.6	(37.9, 100)
Percent of basin with moderate infiltration rate soils (B) (STATSGO)	48.8	(0.5, 100)
Percent of basin with slow infiltration rate soils (C) (STATSGO)	17.4	(0, 74.1)
Percent of basin with very slow infiltration rate soils (D) (STATSGO)	20.7	(0, 76.4)
<b>Flow</b>		
Reach stream flow velocity (m/s)	0.3	(0, 1.1)
Instantaneous discharge (ft <sup>3</sup> /s)	67.7	(0, 481.0)
<b>Study reach habitat</b>		
Wetted width (m)	15.8	(0.5, 93.4)
Bankfull width (m)	12.7	(2.0, 64.0)
Bank vegetative cover (%)	59.0	(13.0, 100)
Percent of substrate = silt (<0.063 mm)	47.8	(0, 100)
Percent of substrate < sand (<2 mm)	34.0	(0, 100)
Percent of substrate > gravel (>2 mm)	36.3	(0, 100)
Substrate embeddedness (%)	58.6	(6.2, 100)
Open canopy angle (°)	100.4	(5.1, 175.7)
Wetted perimeter (m)	16.6	(0.6, 96.8)
Wetted cross-sectional shape ((wetted width/mean depth) exp(mean depth/max. depth))	14.3	(2.0, 46.4)
Max. wetted cross-sectional area (m <sup>2</sup> )	19.7	(0.1, 203.0)
<b>Basin and riparian land cover</b>		
Percent woody vegetation within 100-m buffer	36.5	(0, 85.2)
Percent grassland vegetation within 100-m buffer	17.2	(0, 84.5)
Percent cropland within 100-m buffers	14.9	(0, 95.0)
Percent basin comprised of all agriculture	40.9	(1.6, 99.2)
Percent basin comprised of residential forest	0.1	(0, 1.7)
Percent basin comprised of deciduous forest	12.5	(0, 82.2)
Percent basin comprised of evergreen forest	21.6	(0, 86.5)
Percent basin comprised of shrub	13.4	(0, 86.6)
Percent basin comprised of grasslands	19.8	(0, 93.2)
Percent basin comprised of pasture and hay	5.1	(0, 50.3)
Percent basin comprised of row crops/orchard	11.2	(0, 90.8)
Percent basin comprised of small grains	2.1	(0, 49.8)
Percent basin comprised of emergent wetlands	0.7	(0, 4.6)
<b>Water quality/chemistry</b>		
TP (mg/L)	0.20	(0, 2.69)
TN (mg/L)	1.39	(0.05, 21.20)
Water temperature (°C)	19.0	(1.5, 29.0)

**Table 1** (continued)

Predictor variables	Mean	(Min., Max.)
Specific conductance ( $\mu\text{S}/\text{cm}$ )	372.7	(29.0, 2310.0)
pH	8.1	(7.2, 8.9)
Dissolved oxygen (mg/L)	8.4	(5.2, 12.3)
Dissolved inorganic nitrogen (mg/L)	0.8	(0, 19.4)
Organic carbon (mg/L)	3.8	(0.6, 21.8)
Weather		
Mean annual precipitation (cm)	74.3	(21.9, 146.5)
Invertebrate trophic measures		
Abundance of climber, sprawler, burrower, or swimmer taxa		
Total number of predator, omnivore, collector-gatherers, filtering-collectors, shredder, or scraper taxa		
Total percent of predator, omnivore, collector-gatherers, filtering-collectors, shredder, or scraper taxa		
Total number of predator, omnivore, collector-gatherers, filtering-collectors, shredder, or scraper individuals		
Total percent of predator, omnivore, collector-gatherers, filtering-collectors, shredder, or scraper individuals		

Mean and range values represent results for this study. A detailed discussion of methods used to collect and generate the data is available in Fitzpatrick et al. (1998) and Johnson and Zelt (2005)

*Min.* minimum, *Max.* maximum,  $\text{km}^2$  square kilometers,  $\text{mm}/\text{year}$  millimeters per year, % percent,  $\text{in}/\text{h}$  inches per hour,  $\text{m}$  meters,  $\text{mm}$  millimeters,  $\text{m}/\text{s}$  meters per second,  $\text{ft}^3/\text{s}$  cubic feet per second,  $\text{m}^2$  square meters,  $\text{mg}/\text{L}$  milligrams per liter,  $^{\circ}\text{C}$  degrees centigrade,  $\mu\text{S}/\text{cm}$  microsiemens per centimeter,  $\text{cm}$  centimeters

within 30 and 28 days, respectively. Table 1 includes the water quality variables that were collected or derived and used in the data analysis.

#### Watershed and riparian characteristics

Geographical Information System-derived watershed land cover, soil characteristics, topographic features, runoff estimates, and riparian land cover were calculated for each watershed upstream of the study site. Calculations were made using the National Hydrography Dataset (NHD 100K; US Geological Survey 2003a), the National Land Cover Dataset (NLCD 2001; US Geological Survey 2003b), and digitized orthophoto quarter quadrangles to characterize riparian conditions as described in Johnson and Zelt (2005). Riparian land cover was calculated within 100-m buffers adjacent to the study reaches. The specific watershed and riparian characteristics metrics used in this study are identified in Table 1.

#### Algae and macroinvertebrate sampling

Algal community composition was determined for fine-grained depositional substrate (epipsammic/epipeltic) and coarse-grained erosional substrates (rock >64 mm or wood) within each of

the study reaches immediately prior to the habitat characterization. Coarse substrates were sampled throughout each study reach from five locations and composited into one sample per site. Coarse substrates were sampled by scrapping the biological material off rock or wood from a measured area. The material was transferred into a container by washing the substrate with a known volume of water. The samples were preserved with 5% formalin for identification (Moulton et al. 2002). At five locations within each study reach, depositional substrate was sampled by inverting a 47-mm diameter plastic Petri dish onto the substrate, sliding a spatula under the dish, and rinsing the material into a container and preserving it with 5% formalin for identification. At each location where algae samples were collected, depth, velocity, and substrate characteristics were recorded. Velocity measures were taken as close to the stream bottom as possible. All algal community samples were shipped to the Academy of Natural Science in Philadelphia for algal identification, density, and biovolume calculations (Charles et al. 2002). Algae data for all sites were collected during stable low-flow periods, typically between July and September. Algae and diatoms were identified to the lowest taxonomic level possible. Taxonomic resolution varied among sites and over

time, so the data were modified before data analysis to establish a consistent level of taxonomy across all sites and time. After resolving taxonomic resolution, only those taxa that accounted for more than 0.5% of the density of taxa at two or more sites were retained for further analysis to reduce the influence rare taxa may have on the analysis.

Macroinvertebrate samples were collected from coarse substrate at all sampling sites during low-flow conditions. Coarse substrate was composed of either gravel or cobble typical of riffles or woody debris where sand substrates dominated. A detailed discussion of macroinvertebrate sampling and identification protocol and can be found in Moulton et al. (2000) and Cuffney (2003). Macroinvertebrate behavioral and trophic metrics were created for each site (Table 1). These metrics included the abundance of climber, sprawler, burrower, or swimmer taxa, and the abundance, percent abundance, taxa richness, and percent taxa richness of predators, collector-gatherers, filtering-collectors, scrapers, shredders, or omnivores. Macroinvertebrate behavioral and trophic designations were based on Cuffney (2003).

Major groups of autecological algal attributes or metrics were compared to water chemistry, land use, hydrologic variables, and invertebrate trophic groups. The algal metrics examined in-

cluded indicators of nitrogen metabolism, trophic condition, organic enrichment, dissolved oxygen concentrations, physical conditions, and nutrient optima. The metrics used in this study and the predicted response to increasing nutrients are listed in Table 2. Autecological characterization of algae was based on Prescott (1962), Lowe (1974), Lange-Bertalot (1979), Bahls (1993), van Dam et al. (1994), Potapova and Charles (2007), and Porter et al. (2008).

### Data analysis

To examine the relationships between nutrients and environmental factors collected at watershed and reach scales and algal metrics derived from algal samples collected from two habitat types (fine and coarse substrate), a statistical technique known as random forests (RF) regression was used (Prasad et al. 2006; Cutler et al. 2007). RF is a type of regression tree analysis that does not rely on a priori assumptions about the relationship between response and predictor variables and allows for interactions and nonlinearities among variables. RF produces numerous regression trees from a bootstrapped sample of 70% of the observations for each node or branch in the regression tree. The remaining 30% of the observations are used to evaluate each branch. With this approach, each observation is left out of approximately one

**Table 2** Algal response metrics examined and predicted response to increasing nutrients

Algal metric	Description	Predicted response to increasing nutrients
PS%a	Polysaprobous diatoms (% abundance) <sup>a</sup>	Increase
NH%a	Nitrogen heterotrophic diatoms (% abundance) <sup>a</sup>	Increase
MPT%a	Most pollution-tolerant diatoms (% abundance) <sup>b</sup>	Increase
MPS%a	Most pollution-sensitive diatoms (% abundance) <sup>b</sup>	Decrease
MPS%r	Most pollution-sensitive diatoms (% taxa richness) <sup>b</sup>	Decrease
MA%r	Motile algae (% taxa richness) <sup>c</sup>	Increase
MA%a	Motile algae (% abundance) <sup>c</sup>	Increase
pHr	Alkaliphilus (>7 pH) diatoms (taxa richness) <sup>a</sup>	Increase
TN%a	High TN diatoms—Central/Western Plains, USA (% abundance) <sup>d</sup>	Increase
TP%a	High TP diatoms—Central/Western Plains, USA (% abundance) <sup>d</sup>	Increase
Eur	Eutrophic diatoms—meso-eutrophic to hyper-eutrophic (taxa richness) <sup>a</sup>	Increase

<sup>a</sup>van Dam et al. (1994)

<sup>b</sup>Bahls (1993)

<sup>c</sup>Porter et al. (2008)

<sup>d</sup>Potapova and Charles (2007)

third of the tree, and the final prediction for each observation is obtained by averaging the predictions across all the trees where it was excluded. RF methods have a number of benefits over other regression tree methods. First, RF methods limit overfitting, which enhances its predictive capabilities. Second, the RF can model complex interactions among predictor variables. For example, when predictor variables are highly collinear but good predictors of a response variable, traditional stepwise and criterion-based variable selection procedures will typically retain only one or two of the collinear variables and discard the rest. RF procedures distribute the importance of the collinear predictor variables and guards against eliminating predictor variables that may be ecologically important, but are correlated with other predictors (Cutler et al. 2007). Finally, RF has a robust approach to characterizing the importance of predictor variables. The most important predictors will have the largest decrease in model accuracy when they are removed from the model. The RF method also identifies the percent variance explained for each model. The RF procedure was performed using the R statistical system (Liaw and Wiener 2002; R Development Core Team 2006).

We used two regression techniques to examine the significance of TN or TP thresholds for algal metrics. Linear regression was used to determine if TN and/or TP could predict algal metrics in an expected manner (Table 2) and generate statistics that could be compared with those generated by the second technique known as piecewise regression (Ryan and Porth 2007). Piecewise regression was used to determine if statistically significant thresholds or abrupt changes in the relationship between TN or TP and each of the algal metrics collected from fine and coarse substrate could be identified. Piecewise regression models are “broken-stick” models where two lines are joined at an unknown point called breakpoints. These types of models are effective in modeling abrupt breakpoints or thresholds (Toms and Lesperance 2003). An  $r^2$  value for each piecewise model was produced and this value was compared with that generated for each corresponding simple linear regression model to determine if the piecewise model provided any additional information. For

each threshold value, a 95% confidence interval was calculated as:

$$tCI = t \pm 1.96 (SE)$$

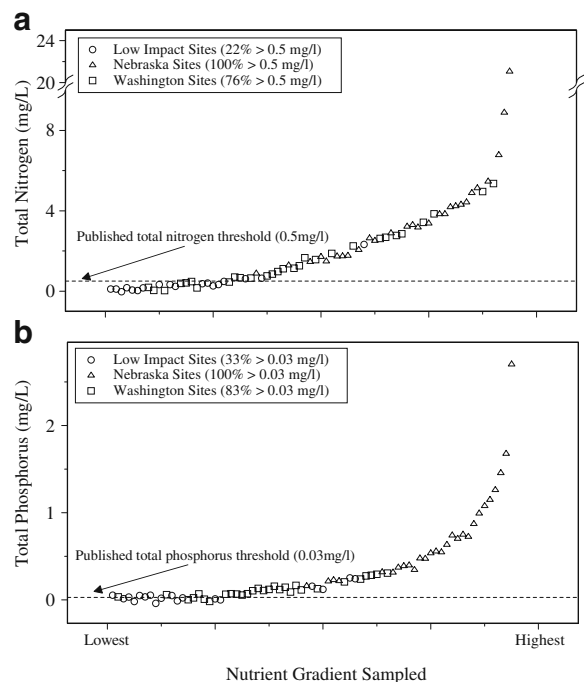
where  $t$  is the threshold and SE is the standard error.

Systat (2004) was used for the linear regression analyses and the nonlinear regression function in Sigma Plot (SPSS 2006) was used for the piecewise regression analyses.

## Results

### Physical and chemical characteristics of streams

Seventy-three sites were evaluated as part of this study, 28 in Washington and Nebraska and 17 low-impact sites. All sites were in the Central and Western Plains’ diatom-based geographic region as determined by Potapova and Charles (2007). The 73 sites had a mean watershed area



**Fig. 2** Sampling sites arranged from lowest to highest mean **a** TN or **b** TP concentrations by location (the site symbols are offset along the y-axis to help identify sampling site locations). TN and TP thresholds were published in Dodds et al. (2002) and Stevenson et al. (2006)



**Table 3** Mean percentage of algae taxa composition by taxonomic group, substrate type, and nutrient concentration and mean percentage of diatom taxa composition by substrate type and nutrient concentration

	Coarse substrate samples		Fine substrate (depositional) samples														
	TP < 0.02 mg/L		TP > 0.24 mg/L		TN < 0.20 mg/L		TN > 2.20 mg/L		TP < 0.02 mg/L		TP > 0.24 mg/L		TN < 0.20 mg/L		TN > 2.20 mg/L		
	% composition	average	% composition	average	% composition	average	% composition	average	% composition	average	% composition	average	% composition	average	% composition	average	
Taxonomic groups—average % composition by nutrient and substrate group																	
Diatoms	82.40	94.17	81.79	93.90	92.88	97.51	93.92	97.63	Green algae	1.42	0.78	1.05	0.80	0.84	0.22	0.83	0.18
Blue–green algae	15.77	3.39	16.82	3.37	5.92	1.88	5.12	1.83	Red algae	0.24	0.62	0.26	0.98	0.00	0.01	0.05	0.05
Yellow–green algae	0.00	0.03	0.00	0.03	0.00	0.00	0.00	0.00	Diatom taxa—average % composition by nutrient and substrate group <sup>a</sup>								
<i>Achnanthydium minutissimum</i>	12.68	0.84	10.67	1.73	10.08	0.85	10.52	2.12	<i>Cocconeis placentula</i>	4.73	2.22	5.29	1.85	3.71	2.48	4.13	0.59
<i>Cyclotella meneghiniana</i>	0.21	2.12	0.06	1.97	0.27	3.38	0.22	3.00	<i>Encyonema</i> sp.	3.59	0.42	3.94	0.60	2.76	0.24	2.33	0.64
<i>Fragilaria vaucheriae</i>	0.55	0.25	1.28	0.36	0.45	0.32	0.84	0.40	<i>Gomphonema</i> sp.	8.98	6.53	7.08	5.65	3.20	2.31	4.17	2.44
<i>Mayamaea atomus</i>	0.01	2.12	0.09	1.52	0.09	0.19	0.21	0.20	<i>Navicula</i> sp.	5.95	23.94	11.23	23.23	9.40	18.64	11.11	18.57
<i>Nitzschia</i> sp.	3.70	29.84	5.31	24.33	8.03	38.04	9.52	33.47	<i>Planorthis</i> sp.	0.68	2.52	1.31	2.17	2.85	3.46	3.86	2.87
<i>Rhoicosphenia abbreviate</i>	0.30	2.57	1.23	2.95	0.34	0.27	0.97	0.49	<i>Sellaphora pupula</i>	0.07	0.62	0.02	0.49	0.52	2.88	0.35	2.38
<i>Synedra ulna</i>	1.95	0.64	1.13	0.86	1.06	0.55	1.01	0.74	Nutrient concentration groups were identified as the 25th and 75th percentiles of the sampled nutrient data. Percent composition was based on cells/area sampled								
<sup>a</sup> Diatom taxa presented were found at over 80% of the study sites																	

of 758.7 km<sup>2</sup> (Table 1). Mean stream velocities for the study reaches collected during habitat and community sampling was 0.30 m/s and ranged from unmeasurable to 1.1 m/s. The mean substrate size was dominated by silts (<0.063 mm), but sands (<2 mm) and gravels (>2 mm) were also found. Across all sites, the mean and range of open canopy measured in the center of the stream channel at each transect (180° equals a completely open canopy) was 100.4° and 5.1–175.7°, respectively (Table 1).

By design, the watershed agricultural land cover ranged from 1.6% to 99.2%. The mean agricultural land cover was 40.9% (Table 1). Crops within 100-m wide buffers adjacent to each study stream ranged from 0% to 95% with a mean of 14.9%. The abundance of woody vegetation within the 100-m buffers had a mean of 36.5%. Although not presented in Table 1, all sites contained <10% urban land cover within their watersheds.

TN ranged from 0.05 mg/L at one of the low-impact sites to a high of 21.21 mg/L at an agricultural site in Nebraska and had a mean of 1.39 mg/L (Table 1). While the study was designed to capture a gradient of nutrient conditions within each agriculturally dominated area, nutrient conditions at most agricultural sites were above a threshold value of 0.5 mg/L for TN published by Dodds et al. (2002) and Stevenson et al. (2006). For the Washington and Nebraska sites sampled, 76% and 100%, respectively, were above the published TN threshold (Fig. 2a). For the low-impact sites, 22% were above the TN threshold (Fig. 2a). TP ranged from less than detection limits to 2.69 mg/L and had a mean of 0.20 mg/L (Table 1). Concentrations at all the Nebraska sites were greater than the Dodds et al. (2002) and Stevenson et al. (2006) published TP threshold of 0.03 mg/L. For the Washington sites, 83% were greater than the published TP threshold. For the low-impact sites, 33% were greater than the TP threshold (Fig. 2b).

#### Algae taxa distribution across sites

At all of the sampling sites for coarse and fine substrate samples, diatoms made up the largest taxonomic group, followed by blue-green, green, red, and yellow-green algae taxa (Table 3).

For both coarse and fine substrate samples, the percentage of taxa identified as diatoms was higher when TP and TN were elevated (Table 3). All of the commonly measured diatom taxa found at ≥80% of the sites were primarily benthic except for *Cyclotella meneghiniana*, which is generally planktonic. *C. meneghiniana* accounted for a relatively small percentage of the diatom taxa found at a site, but accounted for a higher percentage of the taxa when TP and TN were elevated (Table 3). Those diatom taxa that accounted for elevated percentages of observed taxa included *Achnanthydium minutissimum*, *Cocconeis placentula*, *Encyonema* sp., *Gomphonema* sp., *Navicula* sp., *Nitzschia* sp., and *Planorhynchium* sp. (Table 3). All these taxa are incapable of fixing nitrogen and all but *Encyonema* sp. are nonmotile. The average percent composition of diatom taxa was similar between coarse and fine substrate samples (Table 3). However, a number of taxa exhibited higher or lower percent composition under elevated or reduced nutrient levels. For example, samples collected from low-nutrient sites had higher percentages of *A. minutissimum* and *Encyonema* sp. compared to the high-nutrient samples (Table 3). Samples collected from the high-nutrient sites had higher percentages of *C.*

**Table 4** Linear regression results of models relating algal metrics and TN or TP for samples collected from fine- or coarse-grained substrates

Algal metric	TP (mg/L)		TN (mg/L)	
	Fine grain <i>r</i> <sup>2</sup>	Coarse grain <i>r</i> <sup>2</sup>	Fine grain <i>r</i> <sup>2</sup>	Coarse grain <i>r</i> <sup>2</sup>
PS%a	0.45	0.29	0.40	0.16
NH%a	0.29	0.34	0.32	0.37
MPT%a	0.31	0.46	0.30	0.40
MPS%a	0.49	0.51	0.47	0.40
MPS%r	0.51	0.59	0.50	0.51
MA%r	0.32	0.33	0.32	0.33
MA%a	0.13	0.18	0.22	0.20
pHr	0.08	0.19	0.10	0.20
TN%a	0.18	0.24	0.23	0.28
TP%a	0.18	0.26	0.24	0.30
Eur	0.17	0.22	0.18	0.24

Significant parameter estimates and *F* statistics for each model was at the 0.05 level. The slope estimates for each model were consistent with the predicted responses identified in Table 2

**Table 5** Results of RF regression analysis for algal metrics determined from coarse substrate samples and predictor variables

Coarse substrate algal metric	PS %a	NH %a	MPT %a	MPS %a	MPS %r	MA %r	MA %a	pHr	TN %a	TP %a	Eur
Percent variance explained	21.22	62.98	57.57	73.61	63.74	78.47	32.75	52.93	87.97	89.74	65.25
Predictor variables and importance values											
Water quality/chemistry											
TP (mg/L)	18.82	31.46	14.06	30.67	33.90	29.00	19.56	26.74	57.59	71.25	42.01
TN (mg/L)	3.79	12.10	7.24	9.65	19.06	7.63	0.00	11.89	12.56	13.12	13.06
Study reach habitat											
Wetted width (m) (maximum)						31.93					
Substrate embeddedness (%) (mean)						32.83			23.50	27.49	41.58
Percentage of substrate ≤ sand (<2 mm)											
Study basin characteristics											
Soil material < 5 mm (% by weight) (mean)		28.42	27.35	33.97	26.66	31.82	23.50				
Soil material < 5 mm (% by weight) (maximum)								35.80		26.79	
Soil material < 2 mm (% by weight) (mean)	25.47		27.63	32.41							
Soil material < 0.074 mm (% by weight) (mean)											
Rainfall and runoff factor									31.89		40.33
Mean annual basin runoff in 2002 (mm/year)											
Percent silt in basin									28.03		
Percent of basin with slow infiltration rate soils (C) (STATSGO)	21.81										
Basin and riparian land cover											
Percent woody vegetation within 100-m buffer				30.04	23.33						
Percent cropland in study segment within 100-m buffers		23.17									
Percent basin comprised of shrub											
Percent basin comprised of evergreen forest										30.27	
Percent basin comprised of grassland	18.15										
Invertebrate trophic measures											
Percent total richness omnivores			20.56								

The percent variance explained and the three most important predictor variables and/or TN and TP are presented. Variables with the highest importance values will have the largest reduction in model accuracy if removed from the model. Importance values in italics are statistically most important

*PS%a* polysaprobious diatoms(% total abundance), *NH%a* nitrogen heterotrophic diatoms(% total abundance), *MPT%a* most pollution-tolerant diatoms(% total abundance), *MPS%a* most pollution-sensitive diatoms(% total abundance), *MPS%r* most pollution-sensitive diatoms(% total taxa richness), *MA%r* motile algae(% total taxa richness), *MA%a* motile algae(% total abundance), *pHr* alkaliphilus (>7 pH) diatoms(total taxa richness), *TN%a* high TN (>3 mg/L) diatoms—Central/Western Plains, USA (% total abundance), *TP%a* high TP (>0.10 mg/L) diatoms—Central/Western Plains, USA (% total abundance), *Eur* eutrophic diatoms—meso-eutrophic to hyper-eutrophic(total taxa richness)

**Table 6** Results of RF regression analysis for algal metrics determined from fine substrate samples and predictor variables

Fine substrate algal metric	PS%a	NH%a	MPT%a	MPS%a	MPS%r	MA%r	MA%a	pHr	TN%a	TP%a	Eur
Percent variance explained	45.31	44.38	48.00	75.22	72.11	68.91	45.01	54.41	80.80	87.44	59.81
Predictor variables and importance values											
Water quality/chemistry											
TP (mg/L)	32.82	30.05	31.40	34.25	34.77	41.54	9.95	26.23	67.35	66.69	43.47
TN (mg/L)	4.27	14.92	15.22	15.36	16.93	18.17	14.25	12.23	18.23	26.72	12.43
Dissolved inorganic nitrogen (mg/L)		24.55	22.75		28.94						
Study reach habitat											
Wetted width (m) (maximum)							28.78				
Wetted cross-section shape (mean)							32.37				
Substrate embeddedness (%) (mean)											
Percent of substrate $\geq$ gravel (>2 mm)			19.89					31.21	40.31	33.03	29.21
Study basin characteristics											
Soil material < 5 mm (% by weight) (mean)	27.73										
Soil material < 0.074 mm (% by weight) (mean)		23.67		36.71							
Rainfall and runoff factor						22.16		39.60			30.44
Percent of basin with slow infiltration rate soils (C) (STATSGO)						24.85				31.95	
Percent of basin with moderate infiltration rate soils (B) (STATSGO)											
Base flow index				26.99							
Basin and riparian land cover											
Percent woody vegetation within 100-m buffer									28.06		
Percent cropland in study segment within 100-m buffers				31.24	29.86						
Percent basin comprised of row crops/orchard	16.65										
Percent basin comprised of deciduous forest							30.93				
Weather											
Mean annual precipitation (cm)											31.85

The percent variance explained and the three most important predictor variables and/or TN and TP are presented. Those variables with the highest importance values will have the largest reduction in model accuracy if removed from the model. Importance values in italics are statistically most important

*PS%a* polysaprobious diatoms(% total abundance), *NH%a* nitrogen heterotrophic diatoms(% total abundance), *MPT%a* most pollution-tolerant diatoms(% total abundance), *MPS%a* most pollution-sensitive diatoms(% total abundance), *MPS%r* most pollution-sensitive diatoms(% total taxa richness), *MA%r* motile algae(% total taxa richness), *MA%a* motile algae(% total abundance), *pHr* alkaliphilus (>7 pH) diatoms(total taxa richness), *TN%a* high TN (>3 mg/L) diatoms—Central/Western Plains, USA (% total abundance), *TP%a* high TP (>0.10 mg/L) diatoms—Central/Western Plains, USA (% total abundance), *Eur* eutrophic diatoms—meso-eutrophic to hyper-eutrophic(total taxa richness)

*meneghiniana*, *Navicula* sp., and *Nitzschia* sp. when compared to the low-nutrient sites (Table 3). Other studies have shown similar responses to low and high nutrients for these taxa (Bahls 1993; Potapova and Charles 2007).

Relationships between algae metrics and physical and chemical variables

All 11 autecological algal attributes or metrics identified in Table 2 met the predicted responses to increasing nutrients. Although the 11 algal metrics produced statistically significant regression models, the explained variability in the algal metrics ranged from 0.08 for the abundance of fine sediment alkaliphilus diatoms to TP to 0.59 for the percentage of taxa richness of coarse sediment pollution-sensitive diatoms to TP (Table 4).

The RF predictive models also explained a wide range of variability in the 11 algal metrics examined for samples collected from coarse

and fine substrates (Tables 5 and 6, respectively). Models generated for algae collected from coarse sediment generated seven models that explained >60% of the variance in the metrics examined (Table 5). The RF models predicting the percent abundance of high nitrogen and phosphorus diatoms explained 87.970% and 89.74% of the variance, respectively (Table 5). Fifty percent or more of the taxa used to generate the high nitrogen and high phosphorus diatom metrics were the same for all sites examined. For five of the seven models with the highest percent variance explained, the mean TP concentration was the most important predictor variable. For the two models in which TP concentration was not identified as the most important predictor of the algal metrics, watershed soil characteristics were the most important. Although TP was often the most important predictor variable for the RF models generated for algal metrics from coarse substrate, the three to five most important predictor variables for each

**Table 7** TN and TP threshold values by algal metric for fine and coarse grain sediments

Sample substrate and algal metric	TP (mg/L)			TN (mg/L)		
	Threshold	95% confidence interval	Adj. $r^2$	Threshold	95% confidence interval	Adj. $r^2$
<b>Fine grain sample</b>						
Most pollution-tolerant diatoms (% abundance)	0.28	0.11–0.45	0.40	0.86	0.39–1.33	0.31
Alkaliphilus (>7 pH) diatoms (taxa richness)	0.05	0.02–0.07	0.30	a	a	a
Most pollution-sensitive diatoms (% abundance)	0.09	0.03–0.15	0.49	a	a	a
High TN diatoms (% abundance)	0.06	0.04–0.07	0.61	0.77	0.53–1.01	0.44
High TP diatoms (% abundance)	0.06	0.04–0.07	0.71	0.80	0.59–1.01	0.50
Motile algae (% abundance)	0.06	0.02–0.10	0.27	0.99	0–2.65	0.26
Nitrogen heterotrophic diatoms (% abundance)	0.10	0.03–0.17	0.43	1.50	0.76–2.23	0.39
Motile algae (% taxa richness)	0.09	0.02–0.15	0.48	1.49	0.77–2.20	0.39
<b>Coarse grain sample</b>						
Alkaliphilus (>7 pH) diatoms (taxa richness)	0.03	0.02–0.04	0.35	1.25	0.88–1.63	0.30
High TN diatoms (% abundance)	0.07	0.05–0.10	0.71	1.45	0.94–1.97	0.46
High TP diatoms (% abundance)	0.08	0.06–0.09	0.85	1.30	0.93–1.66	0.52
Nitrogen heterotrophic diatoms (% abundance)	0.13	0.08–0.17	0.53	0.59	0.33–0.84	0.40
Motile algae (% abundance)	0.20	0.06–0.33	0.31	a	a	a
Motile algae (% taxa richness)	0.07	0.03–0.10	0.59	1.79	1.08–2.51	0.49

Threshold values, confidence intervals and adjusted  $r^2$  values were determined by piecewise regressions. For each model, all parameter estimates and  $F$  statistics are significant at the 0.05 level

<sup>a</sup> $r^2$  values <0.2 or less than corresponding linear regressions

RF model represented multiple scales. TN never was identified as the most important variable for any of the coarse substrate RF models.

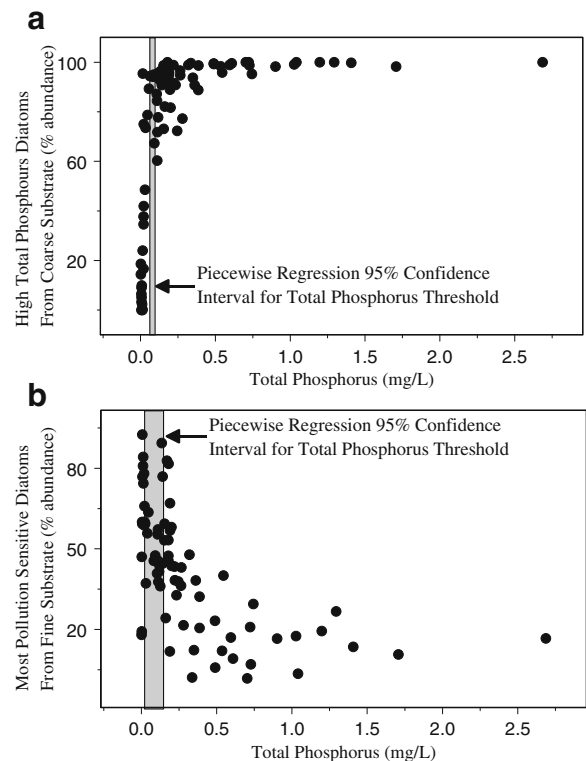
RF models relating the 11 algal metrics collected from fine substrate samples and the 73 physical and chemical predictor variables were similar to those of the coarse substrate samples (Table 6). The fine substrate RF models produced five models that explained >60% of algal metric variance. Like the coarse substrate analysis, predictive models relating the percent abundance of high nitrogen and phosphorus diatoms explained the highest variance at 80.80% and 87.44%, respectively (Table 6). Of the 11 fine sediment algal metric models, TP was identified in eight models as the most important predictor variable (Table 6). Like the RF models for the coarse sediment, TN was not identified as the most important variable in any of fine sediment algal metric models. Excluding TP and TN, the most explanatory predictor variables were rarely similar between coarse and fine sediments.

#### Substrate-specific nutrient thresholds

Piecewise regression was used to identify nitrogen or phosphorus breakpoints or thresholds for each algal metric associated with fine or coarse substrates. In addition to significant parameter estimates, successful piecewise regression models were identified as those models with  $r^2$  values >0.2 (for presentation purposes) and greater than calculated values for the corresponding linear regressions presented in Table 4 (Toms and Lesperance 2003). For both fine and coarse substrates, piecewise regression models were more successful than regression models for TN (Table 7). For fine substrate, eight of the 11 algal metrics produced statistically significant TP models, whereas only six TN fine sediment models were determined to be significant. For coarse substrate, more significant TP models (six) were determined than TN models (five; Table 7). All the significant coarse substrate algal metric models for either TP or TN also were significant for fine substrate samples. However, fine substrate algal metric models produced additional significant models for the percent abundance of the most pollution-tolerant diatoms (TN and TP) and the

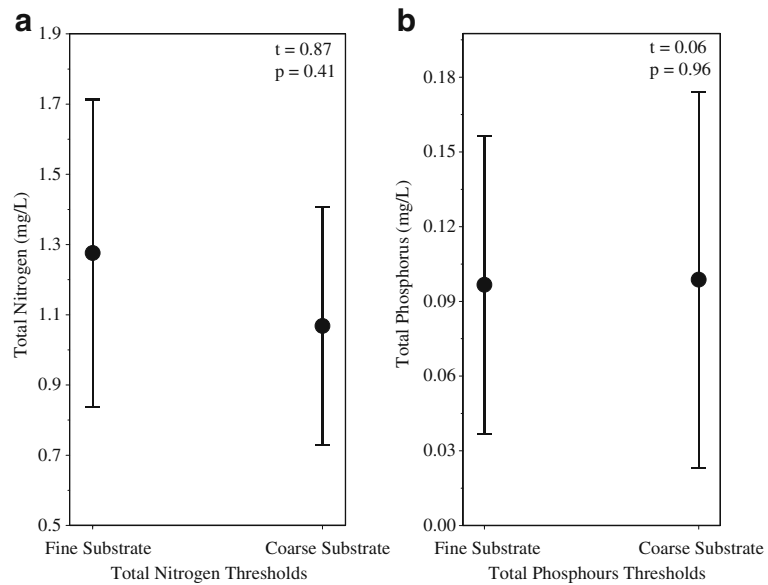
percent abundance of the most pollution-sensitive diatom (TP). For all the piecewise regressions, low levels of either TP or TN produced a large range of potential algal metric scores. The metric values became less variable after the nutrient threshold value was reached (Fig. 3).

The statistically identified threshold values for TP ranged from 0.05 to 0.28 mg/L for fine substrate samples and 0.03 to 0.20 mg/L for coarse substrate samples. For TN, thresholds ranged from 0.77 to 1.50 mg/L for fine sediment samples and 0.59 to 1.79 mg/L for coarse substrate samples (Table 7). A *t* test comparing fine and coarse substrate-generated thresholds determined no significant difference for either TN or TP (Fig. 4).



**Fig. 3** Two examples of piecewise regression plots of total mean phosphorus concentrations against **a** percent abundance of high TP diatoms from coarse substrate and **b** percent abundance of most pollution-sensitive diatoms from fine substrate. The shaded area represents the 95% confidence interval for the phosphorus threshold based on piecewise regression results. Forty-two additional piecewise regression plots were examined (11 TN/fine substrate, 11 TN/coarse substrate, 10 TP/fine substrate, and 10 TP/coarse substrate)

**Fig. 4** Mean and 95% confidence intervals of **a** TN and **b** TP thresholds by substrate type based on piecewise regressions for multiple algal metrics (data used to produce this figure are shown in Table 7)



**Discussion**

Physical, chemical, and taxonomic variation across study sites

As expected, nutrient concentrations, within the agricultural areas, were elevated and generally greater than published threshold values for nitrogen and phosphorus (Dodds et al. 2002; Stevenson et al. 2006). One goal of the study design was to locate sites along a nutrient gradient to examine algal metric responses along such a gradient. However, because of the extensive use of fertilizers and the spreading of animal waste in agricultural areas like the ones in this study, locating low-nutrient sites in agricultural settings for the refinement of nutrient criteria may be difficult. The low-impact sites, a collection of NAWQA sampling sites within the same diatom-based ecoregion as defined by Potapova and Charles (2007), filled in the nutrient gradient and generally had low nutrient levels, but a number of these sites had elevated levels of phosphorus similar to the agricultural sites of Washington and Nebraska. As noted in the “Materials and methods” section, the low-impact sites were identified by the limited amount (<10%) of agriculture or suburban/urban land cover within the watershed of the study site. Although some amount of anthropogenic activ-

ity undoubtedly occurs within these low-impact sites, no effort was made to characterize the proximity of these sources to the sampling sites. A small percentage of anthropogenic activity adjacent to or just upstream of our low-impact sampling sites could have increased nutrient levels. Nevertheless, the low-impact sites provided a gradient of nutrient concentrations that overlapped with some of the agricultural sites. As noted by Potapova and Charles (2007), a broad nutrient gradient will enhance the value of diatoms in the identification of thresholds.

It has been suggested that benthic algae assemblages are similar over very broad geographic areas (Leland 1995; Leland et al. 2001). In the current study, a number of taxa were detected across all sites. These taxa included *C. placentula*, *Gomphonema* sp., *Navicula* sp., *Nitzschia* sp., *Planorhynchium* sp., and *Synedra ulna*. Of these taxa, a number were characterized as nutrient indifferent or had variable responses based on the classification approaches identified in the “Materials and methods” section. For example, *S. ulna* and *C. placentula* have been identified as representative of high nutrients by some (Bahls 1993; van Dam et al. 1994) and low nutrients or indifferent by others (Soininen and Niemela 2002; Potapova and Charles 2007). The high representation of these taxa and the variable trophic designations

for these taxa highlights the potential limitation of using only select diatom taxa metrics derived from different methodologies as an indicator of a stream nutrient status. Given this limitation, this study examined the response of multiple algal taxa compiled in multiple algal metrics to determine if similar responses to changes in physical and chemical variables could be identified.

Although a number of commonly observed diatom taxa were characterized as being indifferent to nutrient concentrations, others were identified as being highly correlated with nutrients. For example, the diatom *C. meneghiniana* has been identified as representative of elevated nutrient levels by a number of studies (Lange-Bertalot 1979; Bahls 1993; van Dam et al. 1994; Potapova and Charles 2007). In this study, the percent composition of *C. meneghiniana* on fine and coarse sediment samples was always higher for samples collected at sites with elevated nutrients when compared to samples with low nutrient concentrations (Table 3). Conversely, the percent composition of the diatom *A. minutissimum*, an indicator of low nutrients, was higher at sites with low nutrient concentrations for both fine and coarse substrate samples (Table 3).

The percent composition of taxonomic groups and select taxa varied more with nutrient concentrations than it did with substrate type (Table 3). These results indicate that sampling either coarse grain substrate in erosional habitats or fine grain substrate in depositional habitats to characterize benthic algal communities may provide similar information. Potapova and Charles (2005) also determined that taxa representation was similar across erosional and depositional sites.

#### Relationships between algae metrics and environmental variables

This study examined the statistical relationship between nutrients, reach to watershed characteristics, and macroinvertebrate trophic characteristics to a suite of algal metrics that often are used in biomonitoring assessments (Stevenson and Bahls 1999; Fore and Grafe 2002; Porter et al. 2008). The metrics examined in this study all responded to changes in nutrients in a statistically significant and expected manner across the wide physical and

geographic range of sites (Tables 2 and 4). For this study, RF regression was used to evaluate the importance of the predictor variables presented in Table 1 in explaining the variation of each of the algal metrics calculated for algal samples collected from fine and coarse sediments. In addition to TP and TN concentrations, predictor variables measured at multiple scales were found to be important predictors of the algal metrics examined. For example, measures of study reach substrate type and embeddedness and study basin soil and runoff characteristics were found to be important predictors of a number of algal metrics for samples collected from both coarse and fine sediments. Our results, like others (Leland 1995; Pan et al. 1999; Potapova and Charles 2003; Soininen et al. 2004; Porter et al. 2008), suggest that multiple environmental factors operating at multiple scales are responsible for structuring benthic algae communities. These environmental factors included nutrients, measures of watershed and instream physical conditions, weather, and macroinvertebrate trophic measures.

Prasad et al. (2006) and Cutler et al. (2007) suggested that RF analyses could often produce very robust statistical relationships. A number of relationships generated for this study explained high percentages of the variability in several of the algal metrics. For example, the RF regressions with the percent abundance of the high nitrogen and phosphorus diatom metrics as the dependent variables explained between 80.80% and 89.74% of the variance. These metrics were developed specifically for the study area (Potapova and Charles 2007). The strong relationship between these two metrics and the physical and chemical explanatory variables examined in this study support the findings of Potapova and Charles (2007), which suggest that metrics developed within a particular area will better reflect the physical and chemical conditions in that study area than in unrelated areas. A surprising finding of the RF model simulation results was the consistently low statistical importance of TN in explaining the variation in the algal metrics, particularly the high nitrogen diatom metric. Porter et al. (2008) also observed poor correlations between similar algae metrics and TN. They suggested that the intensive but periodic use of nitrogen fertilizers might be



responsible for the low correlations. Greater spatial and temporal information on the use of nitrogen fertilizers and instream TN concentrations might improve the relationships between TN and the algal metrics examined. The strong relationship between the high nitrogen diatom metric and TP likely is due to the degree of overlap between taxa characterized as high phosphorus and high nitrogen taxa.

Measures of invertebrate trophic groups were included as predictors of the algal metrics in the RF analysis. Studies have shown that macroinvertebrate herbivores can alter algal biomass and community composition (Steinman et al. 1987; Feminella and Hawkins 1995). However, in this study, trophic groups were a significant predictor in only one of the RF models. While not presented in Tables 5 and 6 due to space limitations, statistical relationships between invertebrate trophic groups and the examined algal metrics were found, but they were less significant than the predictors presented in these tables. While previous studies have found a strong linkage between macroinvertebrate herbivores and algal community structure and density, it has been suggested that generalizations regarding herbivore effects on algae can be tenuous because grazed assemblages can show higher or lower diversity than ungrazed communities (Holomuzki et al. 2006, 2010). Others have shown highly variable grazer effects depending upon spatial heterogeneity in substrate type, nutrient concentrations, and hydrodynamic conditions (Cardinale and Palmer 2002; Peters et al. 2007). It is possible that variable hydrologic and nutrient conditions which were not measured as part of this study could be responsible for the lack of strong relationships between the invertebrate trophic groups and the algal metrics. In addition, the algal metrics selected for this study were selected for their hypothesized relationships with nutrients and as a result may not be significantly influenced by invertebrate trophic groups.

Two algal metrics not included in this study were chlorophyll and ash-free dry weight (AFDW). Previous studies that have tried to relate chlorophyll or AFDW with TN or TP have often determined poor relationships (Munn et al. 1989; Kjeldsen 1994; Dodds et al. 2002;

Porter et al. 2008; but see Biggs and Close 1989; Lohman et al. 1992). Although these two metrics often are examined as part of biomonitoring efforts focusing on nutrient enrichment, a number of studies have determined that these biomass estimates may be more related to hydrologic stability, water clarity, light availability, and the abundance of grazers rather than nutrients (Rosemond et al. 1993; Poff and Ward 1995; Riseng et al. 2004; Biggs et al. 2005). Algal biomass and nutrient relationships may be positive when existing algal biomass is low and negative when biomass is high (Newbold et al. 1982; Mulholland 1996). The dynamic nature of the relationship between chlorophyll and AFDW and nutrient data may make their use in the development of nutrient criteria difficult.

#### Nutrient thresholds

The identification of ecologically relevant response thresholds are of great importance to resource managers. Groffman et al. (2006) defined ecological thresholds as “the point at which there is an abrupt change in an ecosystem quality, property or phenomenon.” Scientists and managers are interested in identifying those drivers responsible for the observed abrupt changes. Evidence of ecological threshold responses has been observed in systems like Florida Bay (Gunderson and Holling 2002) where an abrupt change in water clarity, primary production, and nutrient cycling were tentatively linked to shifts in nutrient inputs and hydrologic changes. The presences of some published thresholds, namely, the urban impervious land cover threshold, however, have been questioned (Morely and Karr 2002). In the case of the urban impervious threshold, Allan (2004) suggested that the response of streams to urbanization is too complex for a single threshold to exist because impervious surfaces are associated with numerous other stressors. Although the authors of this study suggest that the relationship between nutrients and algae is more direct than that of the impervious surface example given by Allan (2004), the RF modeling results of this study did observe a complex relationship between algae, nutrients, and watershed features. The results of this study identified a consistent set of nutrient

threshold values based on numerous algal metrics. The TP threshold values identified in this study (0.03–0.28 mg/L) using algal metrics were similar, but greater than the threshold identified by Dodds et al. (2002) and Stevenson et al. (2006) using chlorophyll *a* (0.03 mg/L). The TN thresholds identified in this study (0.59–1.79 mg/L) were generally higher than the threshold of 0.5 mg/L identified by Dodds et al. (2002) and Stevenson et al. (2006). Although the threshold responses identified in this study used algal metrics rather than chlorophyll or AFDW as the dependent variable, the consistency of the threshold values measured across multiple algal metrics suggest to the authors that the thresholds identified in this study seem to be ecologically relevant. While the threshold values were relatively consistent, the range in the observed values may suggest subtle response difference in some of the algal metrics to nutrient concentrations influenced by additional physical, chemical, or biological variables. In addition, the wide range of algal metric responses measured at low nutrient levels (Fig. 3) and the results of the RF analyses suggest that factors other than nutrients are also affecting algae. A second important finding from this study was the insignificant difference between nutrient threshold values based on algae collected from coarse or fine substrates. These results suggest that managers interested in developing criteria based on algae community data may have to select only one habitat type to sample rather than multiple types.

This study determined that all the algal metrics examined responded to the measured nutrient concentrations in a predictable manner. The statistical importance of nitrogen, phosphorus, or environmental variables in explaining the variation in the algal metrics, however, was highly variable. Phosphorus often was the most statistically important variable to explain the algal metrics and nitrogen typically was one of the least important. As noted earlier, instream nitrogen concentrations based on a limited sampling frequency in agricultural streams with periodic nitrogen fertilizer application are often poorly correlated with diatom metrics (Porter et al. 2008). Two metrics, high nitrogen and phosphorus diatom metrics developed by Potapova and Charles (2007), had the highest percent variance explained by nutrient

and environmental variables. Although phosphorus was the most important predictor for both of these metrics, the high degree of overlap in the taxa used to derive the metrics suggest that these metrics and potentially the other metrics respond similarly to elevated nitrogen or phosphorus.

This study also determined a relatively consistent threshold response of a suite of algal metrics to TN and TP. In addition, these threshold responses were statistically similar for samples collected from coarse and fine substrate. The relationship between these metrics and nuisance growth of filamentous green algae or eutrophic blooms will require additional studies. Nevertheless, the consistent threshold values measured for multiple algal metrics suggests that these threshold values are biologically and management relevant. Although nutrients were typically the most important factors explaining the variation in algal metric variability, physical and chemical features measured at multiple scales also were important. These additional physical and chemical features and their influence on the impact of nutrients on algal communities will need to be used to refine nutrient thresholds in the future.

**Acknowledgments** We thank the National Water Quality Assessment Program of the US Geological Survey for funding this work. We appreciate all of the field and laboratory work of numerous US Geological Survey staff throughout this study.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Allan, J. D. (2004). Landscapes and riverscales: The influence of land use on stream ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, 35, 257–284.
- Bahls, L. L. (1993). *Periphyton bioassessment methods for Montana streams*. Helena: Water Quality Bureau, Department of Health and Environmental Sciences.
- Belore, M. L., Winter, J. G., & Duthie, H. C. (2002). Use of diatoms and macroinvertebrates as bioindicators of water quality in southern Ontario rivers. *Canadian Water Resources Journal*, 27, 457–484.

- Biggs, B. J. F. (1996). Patterns in periphyton of stream. In R. J. Stevenson, M. L. Bothwell, & R. L. Lowe (Eds.), *Algal ecology freshwater benthic ecosystems* (pp. 31–56). New York: Academic.
- Biggs, B. J. F., & Close, M. E. (1989). Periphyton biomass dynamics in gravel bed rivers: The relative effects of flows and nutrients. *Freshwater Biology*, 22, 209–231.
- Biggs, B. J. F., & Kilroy, C. (1994). *Stream periphyton monitoring manual*. Christchurch: The New Zealand Ministry for the Environment.
- Biggs, B. J. F., & Smith, R. A. (2002). Taxonomic richness of stream benthic algae: Effects of flood disturbance and nutrients. *Limnology and Oceanography*, 47, 1175–1186.
- Biggs, B. J. F., Nikora, V. I., & Snelder, T. H. (2005). Linking scales of flow variability to lotic ecosystem structure and function. *River Research and Applications*, 21, 283–298.
- Cardinale, B. J., & Palmer, M. A. (2002). Disturbance moderates biodiversity–ecosystem function relationships: Experimental influence from caddisflies in stream mesocosms. *Ecology*, 83, 1915–1927.
- Charles, D. F., Knowles, C., & Davis, R. S. (Eds.) (2002). *Protocols for the analysis of algal samples collected as part of the U.S. Geological Survey National Water-Quality Assessment Program*. Report No. 02-06. Philadelphia: Patrick Center for Environmental Research, The Academy of Natural Sciences.
- Chételat, J., Pick, F. R., Morin, A., & Hamilton, P. B. (1999). Periphyton biomass and community composition in rivers of different nutrient status. *Canadian Journal of Fisheries and Aquatic Science*, 56, 560–569.
- Cuffney, T. F. (2003). *User's manual for the National Water-Quality Assessment Program Invertebrate Data Analysis System (IDAS) software: Version 3*. US Geological Survey Open-File Report 03-172.
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., et al. (2007). Random forests for classification in ecology. *Ecology*, 88, 2783–2792.
- Dodds, W. K., Smith, V. H., & Lohman, K. (2002). Nitrogen and phosphorous relationships to benthic algal biomass in temperate streams. *Canadian Journal of Fisheries and Aquatic Science*, 59, 865–874.
- Dodds, W. K., Smith, V. H., & Zander, B. (1997). Developing nutrient targets to control benthic chlorophyll levels in streams: A case study of the Clark Fork River. *Water Resources*, 31, 1738–1750.
- Dodds, W. K., & Welch, E. B. (2000). Establishing nutrient criteria in streams. *Journal of the North American Benthological Society*, 19, 1738–1750.
- Feminella, J. W., & Hawkins, C. P. (1995). Interactions between stream herbivores and periphyton: A quantitative analysis of past experiments. *Journal of the North American Benthological Society*, 14, 465–509.
- Fishman, M. J. (1993). *Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—determination of inorganic and organic constituents in water and fluvial sediments*. US Geological Survey Open-File Report 93-125.
- Fitzpatrick, F. A., Waite, I. R., D'Arconte, P. J., Meador, M. R., Maupin, M. A., & Gurtz, M. E. (1998). *Revised methods for characterizing stream habitat in the National Water-Quality Assessment Program*. US Geological Survey Water-Resources Investigations Report, 98-4052.
- Fore, L. S., & Grafe, C. (2002). Using diatoms to assess the biological condition of large rivers in Idaho (U.S.A.). *Freshwater Biology*, 47, 2015–2037.
- Frenzel, S. A., Swanson, R. B., Huntzinger, T. L., Stamer, J. K., Emmons, P. J., & Zelt, R. B. (1998). *Water quality in the Central Nebraska Basins, Nebraska, 1992–95*. US Geological Survey Circular 1163.
- Groffman, P. M., Baron, J. S., Blett, T., Gold, A. J., Goodman, L., Gunderson, L. H., et al. (2006). Ecological thresholds: The key to successful environmental management or an important concept with no practical application? *Ecosystems*, 9, 1–13.
- Gunderson, L., & Holling, C. (Eds.) (2002). *Panarch: Understanding transformations in human and natural systems*. Washington: Island.
- Gurbuz, H., & Kivrak, E. (2002). Use of epilithic diatoms to evaluate water quality in the Karasu River of Turkey. *Journal of Environmental Biology*, 23, 239–246.
- Heinz Center (2002). *State of the nation's ecosystems*. The H. John Heinz III Center for Science, Economics and the Environment.
- Holomuzki, J. R., Feminella, J. W., & Power, M. E. (2010). Biotic interactions in freshwater benthic habitats. *Journal of the North American Benthological Society*, 29, 220–244.
- Holomuzki, J. R., Lowe, R. L., & Ress, A. R. (2006). Comparing herbivory effects of stream macroinvertebrates on microalgal patch structure and recovery. *New Zealand Journal of Marine and Freshwater Research*, 40, 357–367.
- Huntzinger, T. L., & Ellis, M. J. (1993). Central Nebraska river basins, Nebraska. *Water Resources Bulletin*, 29, 533–574.
- Hynes, N. B. (1970). *The ecology of running waters*. Canada: University of Toronto Press.
- Johnson, M. R., & Zelt, R. B. (2005). *Protocols for mapping and characterizing land use/land cover in riparian zones*. US Geological Survey Open-File Report 2005-1302.
- Jones, J. L., & Wagner, R. J. (1995). *Water quality assessment of the central Columbia Plateau in Washington and Idaho: Analysis of available nutrient and pesticide data for ground water 1942–1992*. US Geological Survey Water-Resources Investigations Report 94-4258.
- Kjeldsen, K. (1994). The relationship between phosphorus and peak biomass of benthic algae in small lowland stream. *Verh Internat Verein Limnol*, 25, 1530–1533.
- Lange-Bertalot, H. (1979). Pollution tolerance of diatoms as a criterion for water quality estimation. *Nova Hedwigia*, 64, 285–304.
- 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 Science*, 52, 1108–1129.

- Leland, H. V., Brown, L. R., & Mueller, D. K. (2001). Distribution of algae in the San Joaquin River, California, in relation to nutrient supply, salinity and other environmental factors. *Freshwater Biology*, *46*, 1139–1167.
- 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.
- Liaw, A., & Wiener, M. (2002). Classification and regression by random forest. *R News*, *2/3*, 18–22.
- Lohman, K., Jones, J. R., & Perkins, B. D. (1992). Effects of nutrient enrichment and flood frequency on periphyton biomass in northern Ozark stream. *Canadian Journal of Fisheries and Aquatic Sciences*, *49*, 1198–1205.
- Lowe, R. L. (1974). *Environmental requirements and pollution tolerance of freshwater diatoms*. EPA-670/4-74-005. Cincinnati: Office of Research and Development, National Environmental Research Center, US Environmental Protection Agency.
- Morely, S. A., & Karr, J. R. (2002). Assessing and restoring the health of urban stream in the Puget Sound Basin. *Conservation Biology*, *16*, 1498–1509.
- Moulton, S. R., Carter, J. L., Grotheer, S. A., Cuffney, T. F., & Short, T. M. (2000). *Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—processing, taxonomy, and quality control of benthic macroinvertebrate samples*. US Geological Survey Open-File Report 00-212.
- Moulton, S. R., Kennen, J., Goldstein, R. M., & Hambrook, J. A. (2002). *Revised protocols for sampling algal, invertebrate, and fish communities as part of the National Water-Quality Assessment Program*. US Geological Survey Open-File Report 02-150.
- Mulholland, P. J. (1996). Role of nutrient cycling in streams. In R. J. Stevenson, M. L. Bothwell, & R. L. Lowe (Eds.), *Algal ecology: Freshwater benthic ecosystems* (pp. 609–640). San Diego: Academic.
- Munn, M. D., Black, R. W., & Gruber, S. J. (2002). Response of benthic algae to environmental gradients in an agriculturally dominated landscape. *Journal of the North American Benthological Society*, *21*, 221–237.
- Munn, M. D., Osborne, L. L., & Wiley, M. J. (1989). Factors influencing periphyton growth in agricultural streams of central Illinois. *Hydrobiologia*, *174*, 89–97.
- Newall, P., & Walsh, C. J. (2005). Response of epilithic diatom assemblages to urbanization influences. *Hydrobiologia*, *532*, 53–67.
- Newbold, J. D., O'Neill, R. V., Elwood, J. W., & Van Winkle, W. (1982). Nutrient spiralling in streams: Implications for nutrient limitation and invertebrate activity. *The American Naturalist*, *120*, 628–652.
- 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.
- Patton, C. J., & Gilroy, E. J. (1998). *U.S. Geological Survey nutrient preservation experiment—experimental design, statistical analysis, and interpretation of analytical results*. US Geological Survey Water-Resources Investigations Report 98-4118.
- Patton, C. J., & Kryskalla, J. R. (2003). *Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—evaluation of alkaline persulfate digestion as an alternative to Kjeldahl digestion for determination of total and dissolved nitrogen and phosphorus in water*. US Geological Survey Water-Resources Investigations Report 03-4174.
- Peters, L., Hillebrand, H., & Traunspurger, W. (2007). Spatial variation of grazer effects on epilithic meiofauna and algae. *Journal of the North American Benthological Society*, *26*, 78–91.
- Poff, L. R. (1997). Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society*, *16*, 391–409.
- Poff, N. L., & Ward, J. V. (1995). Herbivory under different flow regimes: A field experiment and test of a model with a benthic stream insect. *Oikos*, *71*, 179–188.
- Porter, S. D., Mueller, D. K., Spahr, N. E., Munn, M. D., & Dubrovsky, N. M. (2008). Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. *Freshwater Biology*, *53*, 1036–1054.
- Potapova, M. G., & Charles, D. F. (2003). Distribution of benthic diatoms in U.S. rivers in relation to conductivity and ionic composition. *Freshwater Biology*, *48*, 1311–1328.
- Potapova, M. G., & Charles, D. F. (2005). Choice of substrate in algae-based water-quality assessment. *Journal of the North American Benthological Society*, *24*, 415–427.
- Potapova, M. G., & Charles, D. F. (2007). Diatom metrics for monitoring eutrophication in rivers of the United States. *Ecological Indicators*, *7*, 48–70.
- Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, *9*, 181–199.
- Prescott, G. W. (1962). *Algae of the Western Great Lakes area*. Dubuque: Brown.
- R Development Core Team (2006). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Retrieved from <http://www.R-project.org>.
- Rantz, S. E. (1982). *Measurement and computation of streamflow: Volume 1. Measurement of stage and discharge*. US Geological Survey Water-Supply Paper, 2175.
- Rinella, J. F., McKenzie, S. W., & Fuhrer, G. J. (1992). *Surface-water quality assessment of the Yakima River basin, Washington: Analysis of available water quality data through 1985 water year*. US Geological Survey Open-File Report 91-453.
- Riseng, C. M., Wiley, M. J., & Stevenson, R. J. (2004). Hydrologic disturbance and nutrient effects on benthic community structure in Midwestern US streams: A covariance structure analysis. *Journal of the North American Benthological Society*, *23*, 309–326.

- Rosemond, A. D., Mulholland, P. J., & Elwood, J. W. (1993). Top-down and bottom-up control of stream periphyton: Effects of nutrients and herbivores. *Ecology*, *74*, 1264–1280.
- Ruddy, B. C., Lorenz, D. L., & Mueller, D. K. (2006). *County-level estimates of nutrient inputs from fertilizer, manure, and atmospheric-deposition sources in the conterminous United States, 1982–2001*. US Geological Survey Scientific Investigations Report 2006-5012. Retrieved from <http://pubs.usgs.gov/sir/2006/5012>.
- Ryan, S. E., & Porth, L. S. (2007). *A tutorial on the piecewise regression approach applied to bedload transport data*. Gen. Tech. Rep. RMRS-GTR-189. Fort Collins: US Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Snyder, E. B., Robinson, C. T., Minshall, G. W., & Rushforth, S. R. (2002). Regional patterns in periphyton accrual and diatom assemblage structure in a heterogeneous nutrient landscape. *Canadian Journal of Fisheries and Aquatic Sciences*, *59*, 564–577.
- Soininen, J., & Niemela, P. (2002). Inferring the phosphorus levels of rivers from benthic diatoms using weighted averaging. *Archiv für Hydrobiologie*, *154*, 1–18.
- Soininen, J., Paavola, R., & Muotka, T. (2004). Benthic diatom communities in boreal streams: Community structure in relation to environmental and spatial gradients. *Ecography*, *27*, 330–342.
- SPSS (2006). *SigmaPlot release 10.0*. Chicago: SPSS, Inc, SigmaPlot.
- Steinman, A. D., McIntire, C. D., Gregory, S. V., Lamberti, G. A., & Ashkenas, L. R. (1987). Effects of herbivore type and density on taxonomic structure and physiognomy of algal assemblages in laboratory streams. *Journal of the North American Benthological Society*, *6*, 175–188.
- Stevenson, R. J. (1997). Resource thresholds and stream ecosystem sustainability. *Journal of the North American Benthological Society*, *16*, 410–424.
- Stevenson, R. J., & Bahls, L. (1999). Chapter six: Periphyton protocols. In M. Barbour, J. Gerritsen, B. D. Snyder, & J. B. Striblind (Eds.), *Rapid bioassessment protocols for use in streams and Wadeable rivers: Periphyton, benthic macroinvertebrates, and fish* (2nd ed., pp. 6.1–6.22). Washington: US Environmental Protection Agency, Office of Water, EPA 841-B-99-002.
- Stevenson, R. J., Rier, S. T., Riseng, C. M., Schultz, R. E., & Wiley, M. J. (2006). Comparing effects of nutrients on algal biomass in streams in two regions with different disturbance regimes and with applications for developing nutrient criteria. *Hydrobiologia*, *56*, 149–165.
- Systat (2004). *Systat 11 Statistics II*. Richmond: Systat Software, Inc.
- Taylor, S. L., Roberts, S. C., Walsh, C. J., & Hatt, B. E. (2004). Catchment urbanization and increased benthic algal biomass in streams: Linking mechanisms to management. *Freshwater Biology*, *49*, 835–851.
- Toms, J. D., & Lesperance, M. L. (2003). Piecewise regression: A tool for identifying ecological thresholds. *Ecology*, *84*, 2034–2041.
- US Environmental Protection Agency (2002). *National water quality inventory: 2000 report*. Washington: US Environmental Protection Agency Report EPA-841-R-02-001.
- US Environmental Protection Agency (2006). *Wadeable streams assessment: A collaborative survey of the nation's streams*. Washington: US Environmental Protection Agency Report EPA 841-B-06-002, Office of Water.
- US Geological Survey (2003a). *Coverages for 1:100k stream tiles, assembled from National Hydrography Dataset*. Reston: US Geological Survey. Retrieved 9 September 2009, from <http://nhd.usgs.gov/index.html>.
- US Geological Survey (2003b). *National Land Cover Database (NLCD 2001) metadata for mapping zones by multilayer downloadable file*. US Geological Survey. Retrieved 9 September 2009, from [http://www.mrlc.gov/nlcd\\_multizone\\_map.php](http://www.mrlc.gov/nlcd_multizone_map.php).
- US Geological Survey (2006). *National field manual for the collection of water-quality data*. US Geological Survey Techniques of Water-Resources Investigations, Handbooks for Water-Resource Investigations.
- van Dam, H., Mertens, A., & Sinkeldam, J. (1994). A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. *Netherlands Journal of Aquatic Ecology*, *28*, 117–133.
- Wolock, D. M. (2003). *Base-flow index grid for the United States*. US Geological Survey Open-File Report 03-263.
- Zelt, R. B., & Jordan, P. R. (1993). *Water-quality assessment of the Central Nebraska Basins—summary of data for recent conditions through 1990*. US Geological Survey Open-File Report 93-422.