Physical Factors Control Phytoplankton Production and Nitrogen Fixation in Eight Texas Reservoirs

Margaret G. Forbes,[1,2]* Robert D. Doyle,[1,2] J. Thad Scott,[1,3] Jacob K. Stanley,[1,2,4] Hui Huang,[2] and Bryan W. Brooks[1,5]

[1]Center for Reservoir and Aquatic Systems Research, Baylor University, One Bear Place #97388, Waco, Texas 76798, USA; [2]Department of Biology, Baylor University, One Bear Place #97388, Waco, Texas 76798, USA; [3]Department of Crop, Soil, and Environmental Sciences, University of Arkansas, 115 Plant Sciences Blvd, Fayetteville, Arkansas 72701, USA; [4]US Army Engineer Research and Development Center, CEERD-EP-R, 3909 Halls Ferry Rd, Vicksburg, Mississippi 39180, USA; [5]Department of Environmental Science, Baylor University, One Bear Place #97266, Waco, Texas 76798, USA

ABSTRACT

We compared regression tree analyses and multiple linear regression models to explore the relative importance of physical factors, land use, and water quality in predicting phytoplankton production and N_2 fixation potentials at 85 locations along riverine to lacustrine gradients within eight southern reservoirs. The regression tree model ($r^2 = 0.73$) revealed that differences in phytoplankton production were primarily a function of water depth. The highest rates of production (mg C m^{-3} h^{-1}) occurred at shallow sites (< 0.9 m), where rates were also related to total phosphorus (TP) levels. At deeper sites, production rates were higher at sites with relative drainage area (RDA, ratio of drainage area to water surface area) below 45, potentially due to longer hydraulic residence times. In contrast, multiple linear regression selected TP, RDA, dissolved phosphorus, and percent developed land as significant model variables ($r^2 = 0.63$). The regression tree

model ($r^2 = 0.67$) revealed that N_2 fixation potentials (mg N m^{-3} h^{-1}) were substantially higher at sites with relatively smaller drainage areas (RDA < 45). Within this subgroup, fixation rates were additionally related to TP values (threshold = 41 μg l^{-1}). The multiple linear regression model ($r^2 = 0.67$) also selected RDA as the primary predictor of N_2 fixation. Regression tree models suggest that nutrient controls (phosphorus) were subordinate to physical factors such as depth and RDA. We concluded that regression tree analysis was well suited to revealing nonlinear trends in data (for example, depth), but yielded large uncertainty estimates when applied to linear data (for example, phosphorus).

Key words: N_2 fixation; primary production; reservoirs; regression trees; total phosphorus; relative drainage area.

INTRODUCTION

The link between nutrient enrichment and increases in planktonic productivity has been well established; however, there is growing evidence that physical factors associated with morphology

and hydrology are the main regulatory controls on phytoplankton (Cuevas and others 2006; Vanni 2006; Arhonditsis and others 2007; Jones and Elliott 2007). The expression of nutrient limitation, although temporally dynamic, may be subordinate to physical factors and often varies along spatial gradients (Smith and Shapiro 1981; Søballe and Kimmel 1987). Additional factors influence rates of primary productivity and N_2 fixation such as temperature (McQueen and Lean 1987; Scott and others 2008), bioavailable nitrogen (Vanderhoef and others 1974; Berman 2001; Scott and others 2008), depth of water column mixing (Levine and Lewis 1987; Sterner 1994), turbulence (Paerl 1985), hydraulic residence time (Dickman 1969; Søballe and Kimmel 1987), length of time of water-column stratification (Paerl 1985), light intensity (Smith and others 1980), sediment quantity and quality (Kimmel and others 1990), sediment oxygen demand (Parr and Mason 2004), and community structure (Sterner 1989; Carpenter and others 2001). Factors affecting planktonic productivity and N_2 fixation have been well described in bioassays and single water bodies. Recent studies have described interactions among variables across a larger spatial scale (Sterner 1994; Patoine and others 2006; Arhonditsis and others 2007; Howarth and others 1988b). Our study examines interaction and hierarchical structure among physical and chemical predictors on a regional scale, which may be the most appropriate scale for evaluating hydrological and ecological processes that are driven by geomorphological and climatological factors (Cuevas and others 2006).

Reservoirs contain a variety of environmental settings in which to examine patterns of carbon and nitrogen fixation potential. Although systems within smaller spatial scales (that is, region or watershed) typically share similar soils and morphology, the hydrodynamics of riverine, transition, and lacustrine zones in reservoirs lead to potentially strong gradients in physical and biogeochemical conditions (Pickett and Harvey 1988; Kimmel and others 1990; Osidele and Beck 2004). Furthermore, land-use patterns in reservoir catchments have been linked to carbon and nitrogen fixation potential through nutrient loadings associated with anthropogenic activities (Arbuckle and Downing 2001; Scott and others 2008). Carbon and nitrogen fixation processes have most often been analyzed by traditional statistical approaches such as linear regression-correlation (McQueen and Lean 1987), multiple linear regression (Patoine and others 2006; Søballe and Kimmel 1987), and nonlinear regression (Smith 1990). Recent studies have employed regression tree analyses (Scott and others 2008) or structural equation modeling (Arhonditsis and others 2007) to evaluate these predominately nonlinear processes. We used regression tree analyses and multiple linear regression to explore spatial patterns of production potential along riverine to lacustrine transects within eight Texas reservoirs. Production potential was obtained from phytoplankton production and N_2 fixation assays on water collected along longitudinal gradients within eight Texas reservoirs. Maximum fixation rates were regressed against physical and limnological predictors such as watershed land use, relative drainage area (RDA), depth, turbidity, and water chemistry.

Classification and regression tree analyses (CART) are often used in ecological studies to define habitat preferred by species or assemblages, particularly in cases where alternative environmental settings can support similar community types (Rejwan and others 1999; Urban 2002; Qian and others 2003; King and others 2005a; King and others 2007). CART has also been used to construct medical diagnosis decision trees because it can handle a large number of diverse predictor variables (nominal or continuous) and the output is relatively simple to interpret (Lewis 2000). The analyses presented here are part of a larger study that examines the relationship between reservoirs zones and water quality. The goal of the study was to provide criteria for delineating locations within reservoirs where conventional ''lake'' water quality standards may not be appropriate. Such locations (for example, reservoir arms) may be fundamentally more conducive to nuisance algal blooms due to location, morphometry, watershed size, or other site characteristics. To evaluate the usefulness of CART in this context, we constructed models with both regression tree analyses (all of our variables were continuous) and conventional multiple linear regression techniques.

Regression tree analyses partition data recursively into subsets that are increasingly homogeneous, providing a tree-like classification that may reveal relationships that are often difficult to reconcile with conventional linear models (Urban 2002). The technique is particularly well suited for resolving nonlinear, hierarchical, and high-order interactions among variables (De'Ath and Fabricius 2000) and for detecting numerical values that lead to ecological changes (Qian and others 2003). General linear models such as analysis of variance can include terms for specific interactions; however, in environmental data sets with many variables the number of potential interactions can

become unmanageable. Regression tree analyses can also reveal structure and hierarchy among interacting variables that are not revealed by traditional linear regression. Regression tree analyses rank continuous variables and therefore do not require data that are normally distributed or that are homoscadistic.

METHODS

We sampled 85 locations within eight reservoirs located along an approximate north-south transect in Texas, USA (Figure 1). To examine spatial variation in phytoplankton production and nitrogen fixation, we selected within each reservoir one generally western arm ''A arm'' and one eastern arm or ''B arm'' (Figure 2). At most reservoirs we sampled three to four sites within each arm, and three or more sites within the main body of the reservoir (C sites). All sampling occurred between July and August 2006 (Table 1), providing a unique opportunity to examine spatial trends on a regional scale under similar climatic conditions. Samples were collected at a depth of 0.3 m. Water depth, temperature, pH, and conductivity were measured in situ. Samples for chemical analyses were stored on ice and transported immediately to the laboratory. Samples for photosynthesis, respiration, and N_2 fixation measurements were stored in lake water and darkness while transported to the laboratory. Turbidity was measured in the laboratory with a Hach 2100N bench top turbidimeter. Chlorophyll a was determined spectrophotometrically after acetone extraction. Total and dissolved nutrients were analyzed with a Lachat Quickchem 8500 Flow Injection Autoanalyzer using standard colorometric techniques (EPA 365.3, 353.2, 365.1, and 353.2).

We determined the phytoplankton production potential of each site by measuring photosynthesis and respiration in light–dark bottle incubations (Fee 1973). Three subsamples were incubated under saturating (375–425 μmol s^{-1} m^{-2}) artificial lighting (6700 K high intensity fluorescent); three subsamples were incubated under low light (33–45 μmol s^{-1} m^{-2}); and three subsamples were incubated in darkness (foil wrapped). Incubations lasted 6–12 h and were maintained at temperature levels observed during reservoir sampling. Respiration was calculated as the decrease in O_2 in dark bottles, whereas maximum gross planktonic production was calculated as the sum of maximum net production and respiration. The average dissolved oxygen change in three replicate clear bottles was converted to carbon equivalents (mg

C m^{-3} h^{-1}) based on an assumed photosynthetic quotient of 1.2. Maximum phytoplankton production potential was calculated as the average of the three highest rates for each site.

Determination of planktonic N_2 fixation potential was conducted using the acetylene reduction method (Flett and others 1975). Nine 30-ml subsamples were drawn into 50-ml Popper® Micromate glass syringes. Five milliliters of acetylene gas were drawn into each syringe and dissolved with gentle agitation. Incubations were conducted simultaneously with phytoplankton production assays, under identical temperature and light conditions. Deionized water was used in multiple syringes under light and dark conditions to serve as method blanks. Following incubations, 15 ml of air were drawn into each syringe and the syringe agitated to establish equilibrium between dissolved and vapor phases. A 100-μl air sample was immediately extracted from each syringe and injected into a Carle® AGC Series gas chromatograph (70°C, helium), equipped with a flame ionization detector and a 1.8-m stainless steel column packed with 80% Porapack N and 20% Porapack Q (80/100 mesh). Multiple 10-ppm ethylene standards were used to calibrate the instrument every 2–3 hours. The acetylene reduction rate was calculated as the average from three maximum replicate clear syringes and converted into a N_2 fixation rate assuming the production of 3 μmols ethylene was equivalent to the fixation of 1 μmol N_2.

Watershed boundaries and land-cover type were determined using ArcGIS v 9.2 with data classes calculated by the 2001 National Land Cover Database (30-m raster coverage). Percent land-cover types were calculated for ''A'' and ''B'' subbasins of reservoir arms sampled as well as for the entire basin of whole reservoirs (''C''). Land covers classified as grassland scrub/shrub, forest, pasture/hay, and barren were aggregated into one class representing undeveloped land. This was necessary due to the tendency of studies encompassing a broad geographic scale to exhibit spatial bias and patchiness of cover types (King and others 2005b). For example, in our study undeveloped areas in eastern reservoirs tended to be dominated by forests whereas undeveloped areas in western reservoirs were dominated by grassland and scrub/shrub. Relative drainage area was calculated as the drainage area divided by the water surface area of the arm or reservoir sampled. For each reservoir, three RDAs were calculated; one for each arm and a third for the entire reservoir. The RDA of an arm was applied to all stations located in that arm whereas the RDA of the entire reservoir was
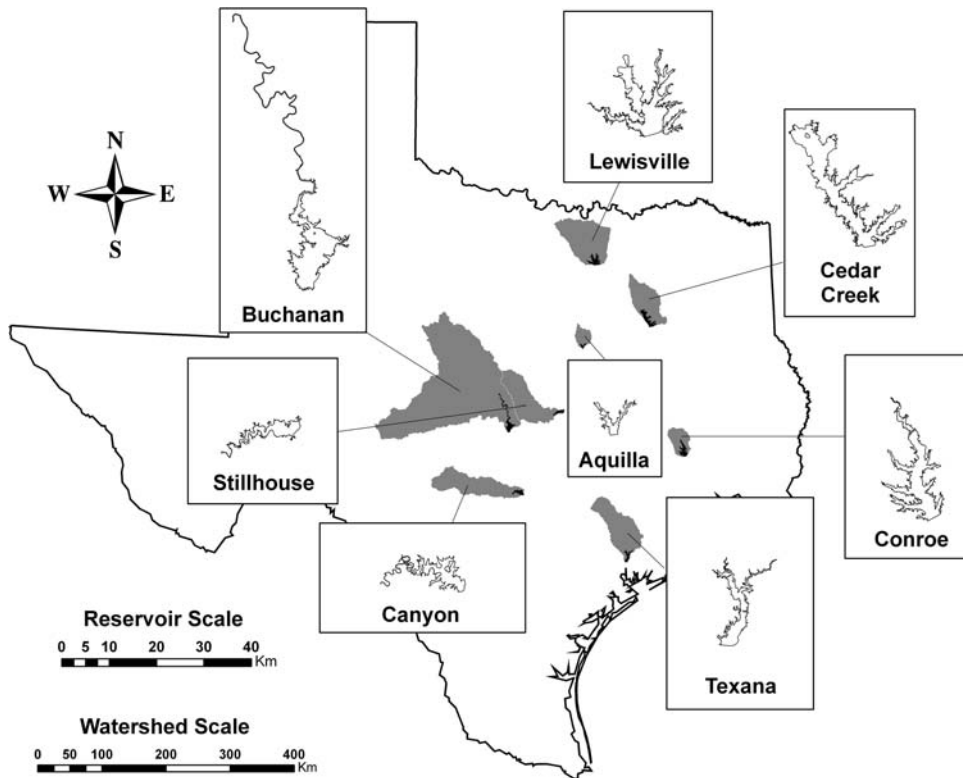
Figure 1. Location of eight Texas reservoirs where carbon and nitrogen fixation were determined for sites located in the main reservoir zone and in one eastern and one western arm (24 sites).

applied to the main reservoir sites. We also calculated the hydraulic retention time (HRT) of each cove and main reservoir based on the nearest gaged streamflow from May 1 through August 29, 2006. Where necessary, gaged data were corrected for differences between the drainage area of the gage and the drainage area of the reservoir or reservoir arm. No adjustments were made for precipitation or evaporation.

Correlations and multiple linear regression (using stepwise forward selection and a significance level for each variable of $P = 0.05$) were conducted using SAS 9.1.3 (SAS Institute, Inc., Cary, NC, USA). Prior to correlation and multiple regression, variables were transformed to meet assumptions of normality. We investigated collinearity by calculating tolerances and variance inflation factors. CART analyses were performed using the RPART library in S-Plus 2000 (Insightful Corp., Seattle, WA, USA). Observations consisted of individual sites ($n = 85$) from eight reservoirs. We required that each split yield a group of at least eight observations and each terminal group of the tree contain at least six observations. We reran both models with reservoir and reservoir arms as a class variable to reveal whether an apparent land use or morphometric effect was driven by observations within a single watershed (King and others 2005a). The recursive nature of regression tree solutions tends to produce models that are sta-

tistically over-fitted to the data used to generate the model, which can compromise the predictive power of the model. To address this, trees are pruned to produce a final model that balances accuracy within the available dataset with robustness to novel data (Urban 2002). Typically, no additional independent dataset is available to test the model, therefore this is accomplished by cross-validation, or V-fold cross-validation, where V usually equals ten, thus yielding ten subgroups of data of similar size (Breiman and others 1984; De'Ath and Fabricius 2000). Cross-validation assures that each variable in the model is sufficiently robust so as to accurately predict observations not included in the model. As each new split (that is, new variable) is added to the model, cross-validation selects 90% of the data, creates the regression tree model, and then uses that model to predict the remaining 10% of the data; this process is repeated using all ten subgroups. The relative error associated with each new model is calculated. This process is repeated, resulting in an over-large tree with an associated overall minimum error (De'Ath and Fabricius 2000). Trees are then pruned to produce the smallest tree with a relative error that also preserves the predictive power of the model. Pruning is at the user's discretion, but typically occurs either at the minimum error or at the smallest tree size that is within one standard error of the minimum error (Urban 2002; Breiman and others 1984). The
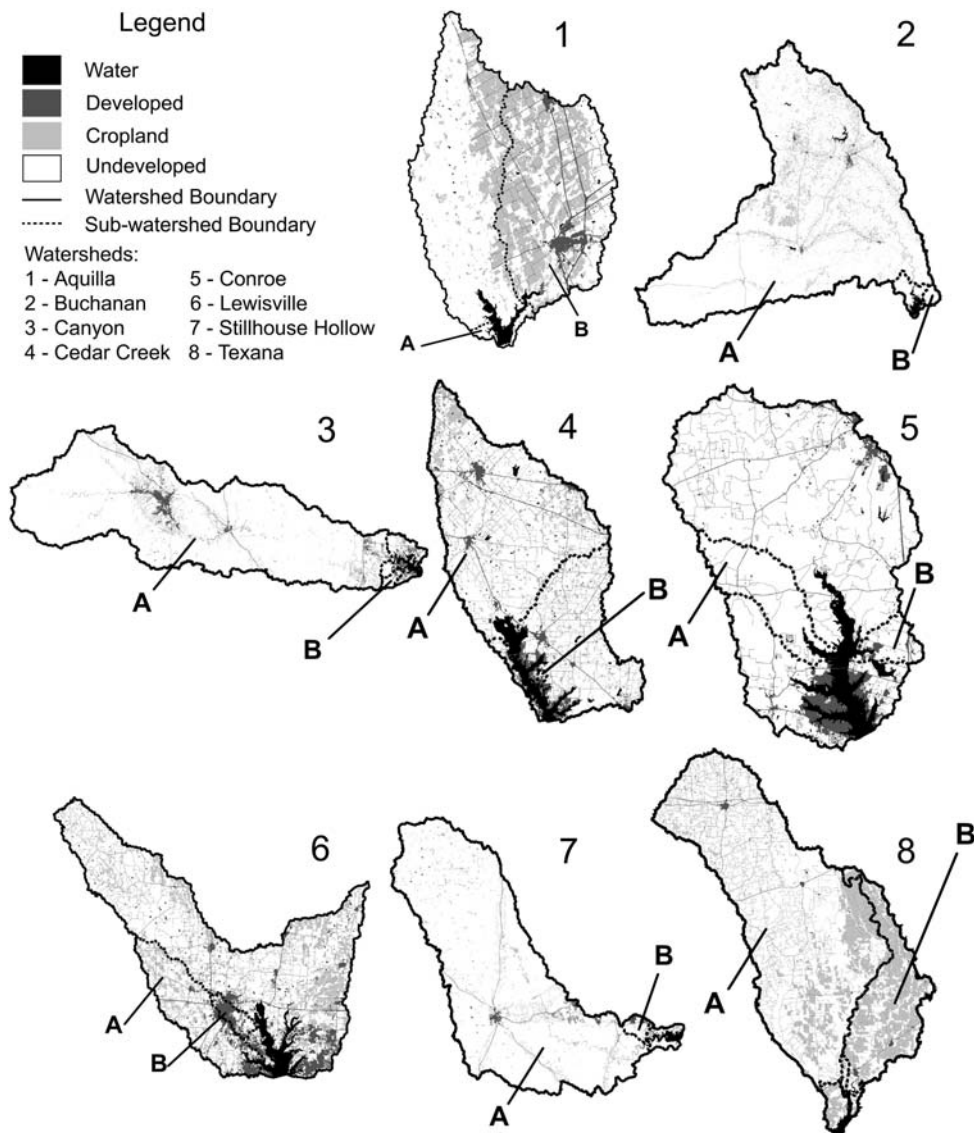
Figure 2. Reservoir (C) and reservoir arm (**A** and **B**) catchment delineations and land use for eight reservoirs. Land use classifications are from National Land Cover Database, 2001.

resulting error associated with regression tree models is the proportion of total sum of squares accounted for by the model compared to the sum of squares of the whole group.

Parameter uncertainty estimates for regression tree predictors were calculated by using an S-Plus changepoint analyses routine on each split of the tree models followed by a bootstrap function (Qian and others 2003). Uncertainty about the value of each split was determined by calculating a 90% confidence interval (CI) from the 1000 bootstrap simulation replicates. The uncertainty was expressed as the 5–95% range from these replicates (Qian and others 2003). We used the approximate χ^2-test to evaluate the statistical significance of each split. The χ^2-test uses the ratio of the deviance reduction to the total deviance prior to the split and

assumes that this ratio is approximately χ^2-test distributed (d.f. = 1) (Venables and Ripley 1994).

RESULTS

Predictor and Response Variables

The observed range of values for response and predictor variables are listed in Table 2. Several variables were strongly correlated (Table 3). For example, water depth was inversely correlated to turbidity ($r = -0.668$), temperatures ($r = -0.568$), and total phosphorus (TP) ($r = -0.549$). Total phosphorus was strongly correlated to N:P ($r = -0.888$); however, total nitrogen was not well correlated to nitrate or N:P. There was a strong association between total nitrogen and chlorophyll ($r = 0.809$), which in turn

Table 1. Characteristics of Reservoirs Sampled for Transitional Zone Study

Label	Reservoir	Latitude	Longitude	Average annual rainfall (cm)	Reservoir area (km²)	Reservoir volume (m³ × 10⁶)	Watershed size (km²)	Maximum depth (m)	Sample date (2006)
1	Aquilla	31°53′N	97°12′W	96.5	13.3	65	645	18.1	7/3
2	Buchanan	30°45′N	98°25′W	67.6	90.4	1080	20357	40.3	8/23
3	Canyon	29°51′N	98°01′W	90.7	33.4	476	3678	38.1	8/16
4	Cedar Creek	32°10′N	96°05′W	98.8	137.7	838	2484	16.2	7/15
5	Conroe	30°21′N	95°33′W	121.4	89.1	531	1085	22.9	7/12
6	Lewisville	33°04′N	96°58′W	87.9	119.8	791	2377	20.4	7/26
7	Stillhouse Hollow	31°01′N	97°32′W	90.4	26.0	291	3376	32.6	8/9
8	Texana	28°53′N	96°33′W	101.9	39.4	210	3540	17.7	8/2

was well correlated to both production potential ($r = 0.918$) and N_2 fixation potential ($r = 0.793$). Hydraulic retention time was inversely correlated to RDA ($r = -0.603$). Some of these correlated variables, such as chlorophyll, total nitrogen, and turbidity were eventually eliminated as predictors as later discussed.

Land-use cover varied widely among the eight reservoirs and between reservoir arms. Three reservoirs, Buchanan, Canyon, and Stillhouse Hollow, were largely undeveloped, with watershed land use primarily comprised of grass/shrub and forest. Proportions of developed land were highest in Lewisville-B and Cedar Creek-B subbasins (0.57 and 0.43, respectively), which are both located in large metropolitan areas. In addition, the tributary for Lewisville-B receives up to 21 million gallons per day of treated municipal wastewater, which is discharged approximately 3 km upstream of this study's sampling locations. Cedar Creek-B is a small cove with high-density residential development along its shoreline. Moderately developed subbasins included Stillhouse-B (0.20) and Lewisville-A (0.16). Proportion of cropped land was highest in Texana-B (0.56) and Aquilla-B (0.43) subwatersheds. Western subbasins of Lewisville and Texana also had considerable areas of cultivated crops, with the result that Aquilla, Texana, and Lewisville whole reservoir watersheds had the highest densities of cultivated crop areas.

Multiple Linear Regression Models

Linear regression models for both phytoplankton production and N_2 fixation potentials each contained four parameters (Table 4). None of the parameters had tolerance values less than 0.1 and thus no eliminations were made based solely on collinearity. Phytoplankton production potentials were related to TP (Figure 3), dissolved (soluble reactive) phosphorus, RDA, and percent of developed land. The overall r^2 for the phytoplankton production model was 0.63 (adjusted $r^2 = 0.62$). The multiple regression model for N_2 fixation also included RDA (Figure 3) and dissolved phosphorus, with RDA providing the greatest contribution to the model ($r^2 = 0.43$). N_2 fixation was also associated with total nitrogen and percent of undeveloped land. The overall adjusted r^2 for the N_2 fixation model was 0.67 (adjusted $r^2 = 0.65$).

Regression Tree Models

The pruned regression tree model for phytoplankton production (Figure 4) included three variables: depth, RDA, and TP. Depth and RDA accounted for

Table 2. Predictor and Response Variables, Minima, Maxima, and Medians of Values

	Minima–maxima	Median
Response variables		
Nitrogen fixation potential (mg N m^{-3} h^{-1})	0–11.7	0.7
Phytoplankton production potential (mg C m^{-3} h^{-1})	6–688	134
Predictor variables		
Land use (percent cover)		
Developed	0–58	5
Undeveloped	38–99	86
Cultivated crop	0–57	3
Depth (m)	0.3–23	4.0
Relative drainage area (unitless)	4–556	76
Hydraulic retention time (year)	0.1–69	1.6
Temperature (C)	24–36	30[a]
Turbidity (NTU)	0.5–416	8.3
Total phosphorus (µg l^{-1})	3.6–901	38
Dissolved phosphorus (µg l^{-1})	1.4–270	10
Total nitrogen (µg l^{-1})	117–2030	570
Nitrate+nitrite nitrogen (µg l^{-1})	2.1–145	4.5
Total nitrogen:Total phosphorus (molar)	3.8–340	31
Chlorophyll (µg l^{-1})	1.2–150	24

[a]*Mean and median temperature.*

the largest share of variability in phytoplankton production among sampling sites (partial $r^2 = 0.43$ and 0.16, respectively). Parameter values, uncertainty estimates, and statistical significance *P*-values are summarized in Table 5. The depth threshold was 0.9 m (90% CI = 0.9–1.7, $P < 0.001$). Shallow sites (<0.9 m) had higher production rates than deeper sites and these shallow sites were further split into two groups based on a TP threshold (206 µg l^{-1}, partial $r^2 = 0.14$). The uncertainty range for the TP threshold was large (90% CI = 65–452 µg l^{-1}). The highest mean (±SD) production rates (581 ± 144 mg C m^{-3} h^{-1}) occurred in this group of shallow sites with high TP concentrations. At this split, the regression tree model listed dissolved phosphorus (split threshold of 21 µg PL^{-1}) as an alternative variable with identical model improvement. In the group of 73 sites deeper than 0.9 m, a further split occurred at RDA of 45 (CI = 42–56, partial $r^2 = 0.16$). At sites with larger RDAs, phytoplankton production rates were the lowest observed in the study (mean 86 ± 49 mg C m^{-3} h^{-1}), whereas the group with RDA below 45 had mean primary productivity that was nearly three times higher. Thus, production rates were lowest at deeper reservoir sites with larger RDAs, and nutrient controls could be validated only at shallow sites. The regression tree model implied a hierarchical structure among variables where depth separates the sites into different "habitats." Primary production at very shallow sites

was further influenced by phosphorus, whereas production at deep sites was associated with relative watershed size (RDA). The relative error (r^2) explained by the regression tree was 0.73. Splits that were pruned from the tree included two additional TP splits ($r^2 = 0.06$ and 0.02) and percent crop land ($r^2 = 0.02$).

Examination and subsequent elimination of competing explanatory variables can lead to a more complete understanding of the predictor variables and better, simpler models (De'Ath and Fabricius 2000). To address competing (correlated) variables, we eliminated two predictors from the initial regression tree analysis for phytoplankton production: total nitrogen and turbidity. When included in the model, total nitrogen controlled most of the splits. The strong association between phytoplankton production rates, chlorophyll, and total nitrogen (Figure 5A) suggests that total nitrogen may be largely derived from phytoplankton. Following elimination of total nitrogen, turbidity was initially selected over depth for the first split. However, turbidity and depth were closely correlated ($r = -0.668$). Furthermore, the positive correlation between turbidity and phytoplankton production (and chlorophyll, Figure 5B) suggests that a sizeable portion of the turbidity may also be associated with algal cells. Turbidity is also associated with light limitation, which inhibits phytoplankton production rates. We therefore eliminated turbidity

Table 3. Pearson Correlation Coefficients for Predictor and Response Variables

Variable	DEV	UNDEV	CROP	Log depth	Log RDA	Log HRT	Log TP	Log DP	1/Log NO$_3$	Log NP	Log Turb	Temp	Log TN	Log Chl	Log PP	Log Nfix
DEV	1	−0.661	−0.009	−0.273	−0.225	0.168	0.329	0.194	−0.043	−0.228	0.374	0.118	0.328	0.319	0.410	0.422
UNDEV		1	−0.730	0.186	0.025	0.144	−0.531	−0.538	0.214	0.534	−0.445	0.063	−0.272	−0.276	−0.298	−0.130
CROP			1	0.004	0.160	−0.355	0.426	0.520	−0.268	−0.505	0.279	−0.235	0.099	0.137	0.075	−0.150
Log Depth				1	0.136	0.022	−0.549	−0.145	0.028	0.415	**−0.668**	−0.568	−0.496	−0.494	−0.559	−0.229
Log RDA					1	**−0.603**	−0.045	0.168	−0.363	−0.134	−0.032	−0.111	−0.304	−0.403	−0.464	**−0.660**
Log HRT						1	−0.249	−0.308	0.420	0.289	−0.216	−0.105	−0.066	0.025	0.105	0.393
Log TP							1	**0.692**	−0.339	**−0.888**	**0.757**	0.282	**0.696**	**0.628**	**0.608**	0.148
Log DP								1	−0.449	**−0.688**	0.471	0.086	0.366	0.137	0.161	−0.217
1/Log NO$_3$									1	0.391	−0.277	0.120	−0.095	0.050	0.035	0.357
Log NP										1	**−0.695**	−0.120	−0.287	−0.319	−0.332	0.127
Log Turb											1	0.241	0.490	0.532	0.532	0.146
Temp												1	0.399	0.298	0.332	0.137
Log TN													1	**0.809**	**0.747**	0.520
Log Chl														1	**0.918**	**0.793**
Log PP															1	**0.778**
Log Nfix																1

Land uses (DEV, UNDEV, and CROP) were arcsine square root transformed and most environmental variables were log-transformed to meet assumptions of normality. Correlations greater than 0.6 are shown in bold. PP refers to phytoplankton productivity.

Table 4. Multiple Linear Regression Statistics Describing Dependence of Phytoplankton Production and Nitrogen Fixation Potentials on Predictor Variables

Variable	Parameter estimate (standard error)	$P > F$	Partial r^2 (Adjusted r^2)
Phytoplankton production model			
Log TP	0.55 (0.07)	<0.0001	0.356
Log RDA	−0.23 (0.05)	<0.0001	0.192
Log DP	−0.24 (0.06)	0.0005	0.064
DEV	0.83 (0.39)	0.0356	0.022
Model summary		<0.0001	0.634 (0.615)
Nitrogen fixation model			
Log RDA	−0.84 (0.15)	<0.0001	0.428
Log TN	2.09 (0.36)	<0.0001	0.122
Log DP	−0.78 (0.16)	0.0006	0.072
UNDEV	−2.54 (0.86)	0.0043	0.044
Model summary		<0.0001	0.666 (0.646)

Variables that did not meet the 0.05 significance level were not included in the model.

as a predictor because we could not resolve its interdependence on depth and algal cell density, or its effect on light attenuation.

RDA was the most important predictor in the N_2 fixation tree model (partial $r^2 = 0.44$, Figure 6). The mean RDA threshold (45) was the same for both N_2 fixation and phytoplankton production. Sites with RDA above 45 had the lowest fixation rates (mean 0.299 ± 0.47 mg N m^{-3} h^{-1}) and this group could not be split further. Sites with RDA below 45 had the highest mean N_2 fixation rates (3.94 ± 3.22 mg N m^{-3} h^{-1}) and were further split according to their TP concentrations (threshold = 40 µg l^{-1}). Uncertainty analysis indicated that the 90% CI for TP ranged from 39 to 156 µg l^{-1}. Sites with higher TP levels had mean N_2 fixation rates over three times higher than sites with lower TP (partial $r^2 = 0.23$). Two additional splits on the high TP branch were pruned from the model. The first was based on N:P (threshold value = 28 molar, partial $r^2 = 0.04$) and an additional split of the low N:P group by nitrate-nitrogen (threshold value of 4 µg l^{-1}, partial $r^2 = 0.01$). The relative error explained by the pruned regression tree was 0.67.

DISCUSSION

In this study, we evaluated physical and biochemical factors that potentially support the growth of phytoplankton and N-fixing cyanobacteria at a regional spatial scale. We focused on identifying variables associated with longitudinal gradients (that is, riverine to lacustrine) that predict spatial patterns of phytoplankton blooms within and

among eight reservoirs. We found that regression tree analyses predicted production potential as a function of depth, with TP acting as a secondary
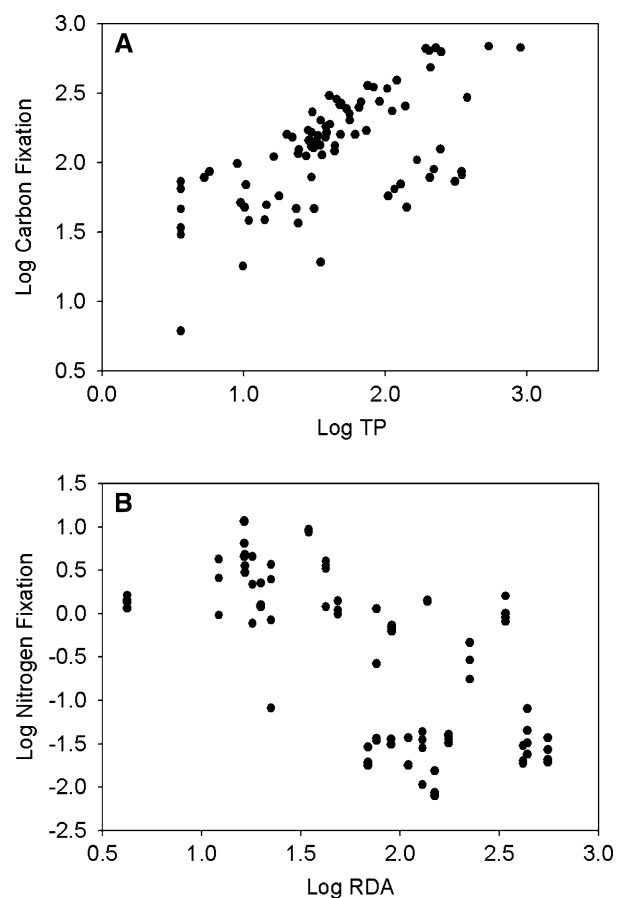
Figure 3. Data of best single predictors from multiple linear regression of (**A**) phytoplankton production potential and (**B**) nitrogen fixation.
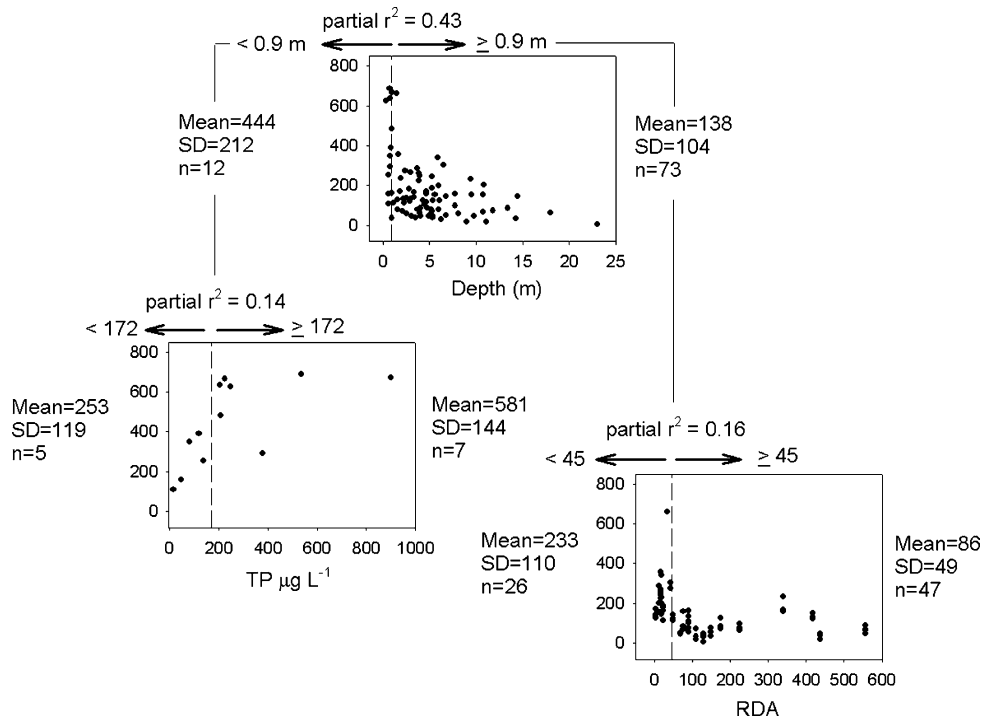
Figure 4. Results from CART analysis of phytoplankton production potential (mg C m^{-3} h^{-1}). Scatterplots illustrate the relationship between production potential and selected parameters at each level of the tree. The vertical *dashed line* in each plot identifies the value of the predictor (x) that best explained variation in phytoplankton production (y). Threshold values of predictors are shown to the left and right of each split above each scatterplot. Variance explained (r^2) for predictors is shown above each split. Means, standard deviation (SD), and number of samples (n) for each subset of data are shown to the left and right of each split. The total variability explained by this CART model was 73%.

Table 5. Regression Tree Threshold Values, Uncertainty Estimates, and P—Values for Phytoplankton Production and Nitrogen Fixation Models

| Variable	Threshold (90% CI)	$P > \chi^2$	Partial r^2
Phytoplankton production model			
Depth (m)	0.9 (0.9–1.7)	<0.0001	0.43
RDA	45 (42–56)	<0.0001	0.16
TP (μg l^{-1})	172 (65–457)	<0.0001	0.14
	Model summary		0.73
Nitrogen fixation model			
RDA	45 (38–45)	<0.0001	0.44
TP (μg l^{-1})	40.5 (39–156)	<0.0001	0.23
	Model summary		0.67

Uncertainty estimates were calculated using a bootstrap method (Qian and others 2003) and P values were based on deviance reduction assuming a χ^2 distribution.

control at shallow sites and RDA acting as a secondary control at deeper sites. However multiple linear regression did not identify depth as an important factor, most likely due to the strong nonlinear relationship between depth and phytoplankton production (Figure 4). Instead, multiple linear regression identified TP as its most significant predictor. It has been well established that

increased levels of phosphorus increase freshwater phytoplankton production in more or less linear trends (Schindler 1978; Smith and Shapiro 1981; Carpenter and others 1998a). At the same time, temporal (Huppert and others 2002) and spatial (Smith and Shapiro 1981) studies have indicated that there are threshold levels of phosphorus below which blooms do not generally occur, and that
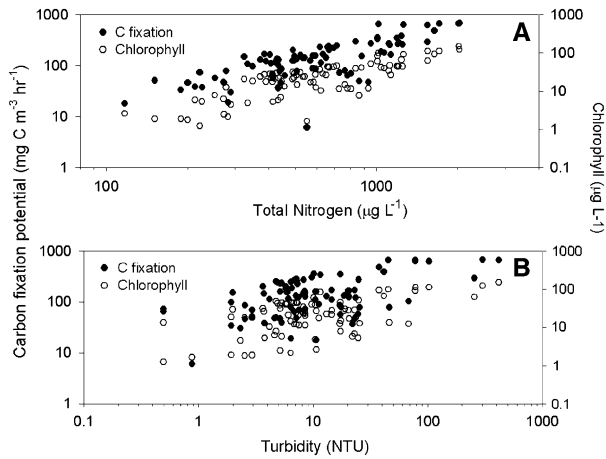
Figure 5. Correlated water quality data from 85 sites in 8 reservoirs. (**A**) Relationship between phytoplankton production potential and total nitrogen (*solid circles*) and chlorophyll and total nitrogen (*open circles*). (**B**) Relationship between phytoplankton production potential and turbidity (*solid circles*) and chlorophyll and turbidity (*open circles*).
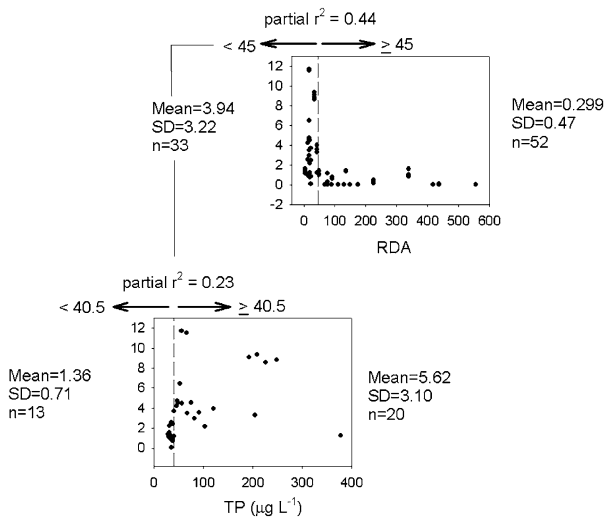
Figure 6. Results from CART analyses of nitrogen fixation (mg N m^{-3} h^{-1}). See Figure 4 for details. The total variability explained by this CART model was 67%.

combinations of light limitation, temperature, and nutrient limitation simultaneously control actual primary production (Sterner 1994).

Results of our study are only partially consistent with trends predicted by heuristic theories of reservoir zonation. These theories, developed by Kimmel and others (1990), Thornton, Kennedy and Walker (1990), and others (see Thornton and others 1990) state that reservoirs can be divided into riverine, transitional, and lacustrine zones.

Although the boundaries between these zones are poorly defined and temporally dynamic, riverine areas are generally characterized by shallow, narrow, well-mixed water columns, with velocities sufficient to transport significant quantities of finer suspended particles such as silts, clays, and organic particles (Gordon and Behel 1985). Increased sedimentation occurs in the transition zone with corresponding increases in light levels (Kimmel and others 1990). The deepest sites occur in the lacustrine zone which is most similar to natural lake systems with low inorganic particulates, higher light penetration, stratification, and increased nutrient limitation (Thornton 1990). Volumetric phytoplankton biomass and primary productivity per unit volume is predicted to be high in the riverine zone, higher still in the transitional zone, and lowest in the lacustrine zone (Kimmel and others 1990). Note that these trends are not necessarily true for integrated areal production rates.

In contrast, we found that volumetric rates were highest in very shallow waters, followed by transitional zone sites and lacustrine sites. Shallow sites were more turbid, had higher phosphorus levels, and warmer temperatures than deeper sites, which is generally consistent with reservoir zonation theory (Kimmel and others 1990). Thus, despite potential light limitation associated with higher turbidities, our production rates were highest at shallow (riverine) sites. Factors contributing to higher phosphorus levels in shallow areas include proximity to nutrient sources (Kimmel and others 1990) and phosphorus releases associated with sediment resuspension processes such as wind-induced, turbulent mixing, and molecular diffusion (Thomas and Schallenberg 2007). Interactions between depth and TP levels in reservoir systems have been observed by others. For example, Sterner (1994) documented higher incidences of nutrient limitation in shallow (for example, riverine and transitional) areas of a Texas reservoir. Søballe and Kimmel (1987) found TP levels in rivers (mean depth 3.2 m) were over twice those found in impoundments (mean depth 8.9 m); however, algal abundance in rivers, impoundments, and natural lakes had similar responses to TP and TN. In fact, their parameter estimate for log TP was 0.55 for rivers, and 0.48 for both natural lakes and impoundments. This is nearly identical to the parameter estimate from our linear regression model for phytoplankton production (0.55 ± 0.07). Thus our sites exhibited good linear response to TP (partial $r^2 = 0.359$) along the longitudinal gradient from riverine to lacustrine sites. However, Søballe and Kimmel found that both depth and

residence times were important copredictors in their multiple linear regression models. Furthermore, nonlinear responses to residence time were observed that led the authors to group systems by residence times ($<$75 and $>$120 days) and conclude that physical controls of depth and residence time had an hierarchical effect on nutrient relationships across broad spatial scales. Furthermore, the relative importance of these factors may change along environmental gradients such as depth or riverine–lacustrine transects. Thus, we concluded that the regression tree analyses, which emphasized the role of phosphorus at shallow sites, provides a more useful model for identifying *where* in these reservoirs algal blooms are likely to occur.

Although higher phosphorus levels and temperatures positively affect production rates, turbidity is generally associated with lower growth rates due to reduced illumination (Søballe and Kimmel 1987; Grobbelaar 1989). Smayda (1970) argued that higher turbidity may reduce sinking losses, enhance nutrient acquisition, and increase light exposure of suspended algae. In shallow lakes with clay turbidities, the relative importance of these factors is not completely understood (Lind and others 1992). For example, in Lake Chapalla, Mexico, phytoplankton was most productive per unit volume at the shallowest and most turbid station, whereas the least turbid station, with a deeper circulating water column, had the lowest volume-based production (Lind and others 1992). Lind found that clay turbidity differentially affected vertical attenuation coefficients for different spectra (red light was 4.6 m^{-1} whereas green light was 6.2 m^{-1}) and thus may reduce photoinhibition in shallow waters. Clay turbidity has also been associated with reduced effectiveness of fish predation and weakening of the link between zooplankton and phytoplankton in Scandinavian lakes (Horpilla and Liljendahl-Nurminen 2005). Clays and similar suspended sediments have large surface areas that often carry sorbed phosphates and other compounds. In fact, reservoir food web models identified the tight coupling between suspended sediments and phosphorus loading as one of the most important factors affecting attainment of trophic goals (Osidele and Beck 2004). Reservoirs in our study were primarily located within the Blackland Prairie Ecoregion which is characterized by iron- and aluminum-rich clays. Turbidity-depth gradients in these types of reservoirs may be reversed in other regions (for example, Scandinavia) where turbidities are derived primary from autochthonous sources and reach maxima in deeper waters (Horpilla and Liljendahl-Nurminen 2005). Thus, the relationships between depth, turbidity, phosphorus, and phytoplankton production rates observed in our study may not be valid for other ecoregions. Although these trends vary for areal-based production, poor water quality associated with nuisance blooms are based on volumetric measurements (for example, low dissolved oxygen concentrations). Thus, depth appears to be an effective aggregate variable for identifying reservoir zones where elevated algal production and related poor water quality are most likely to occur.

At deeper ($>$1 m) sites, phytoplankton production potentials were best predicted by RDA. Relative watershed size has been used by wetland scientists as a correlate of sediment and nutrient retention, flushing, upstream erosion, and other hydrological processes (Adamus and others 1983). In hydrogeomorphic (HGM) wetland functional assessment, normalized watershed:wetland ratios have been used to estimate runoff processes within regional wetland types, and to make inferences about ecological and hydrological functions. For example, in the HGM model for U.S. prairie potholes (Gilbert and others 2006), ratios of catchment:wetland area are used to predict whether a wetland is likely to have groundwater recharge. An important aspect of HGM is that all variables are normalized to reference wetlands of the regional type; thus, what is important is the relative value of the predictor. Thus, the threshold value selected by regression tree analyses would likely be different in regions with different geomorphology and climate.

In limnological studies, several authors have evaluated drainage area (that is, watershed size) as a predictor of ecological processes. In an analyses of hundreds of sites across the United States, drainage area and water residence time were correlated to algal biomass ($r = 0.7$ for both) in rivers, lakes, and reservoirs (Søballe and Kimmel 1987). Relative watershed sizes are larger for reservoirs than for natural lakes (Thornton and others 1981), thus potentially enhancing the role of physical processes such as nutrient, sediment, and hydrologic loading (Vanni and others 2006; Maberly and others 2002). Furthermore, there is evidence that, on a regional scale, watershed size is a good predictor of peak runoff volumes. Harmel and others (2006) found a strong linear relationship between log of watershed area and peak discharge data for 17 watersheds in the Blackland Prairie Ecoregion with little scatter for 2, 5, 10, 25, 50, and 100-year return intervals. Others have concluded that faster flushing times tend to reduce phosphorus levels and thus reduce phytoplankton levels (Vollenveider and Kerekes 1980; Maberly and others 2002). Despite these

findings, the use of RDA as a proxy for hydrologic forcing is experimental and would be expected to vary geographically. In addition, relationships between RDA and processes such as reservoir nutrient loading and flushing rates are largely theoretical.

The impacts of increased flushing rates on phytoplankton production remain complex and often contradictory. Hydrological forcing has been associated with altered water residence times, stratification patterns, and chemical and nutrients loads that differentially affect phytoplankton growth (Arhonditsis and others 2007). Theoretically, high flushing rates would be expected to result in higher nutrient levels because sites are more closely linked to terrestrial nutrient sources (Søballe and Kimmel 1987). However, high flushing rates are also linked to higher turbidities and cell wash-out (Dickman 1969), both of which may decrease phytoplankton abundance. Combined effects of these interacting processes are probably dependent on location (Vanni and others 2006; Cuevas and others 2006) and other factors. For example, when flushing rates do not exceed the mean doubling time of the phytoplankton assemblage, increased inflows may enhance phytoplankton productivity by increasing nutrient availability (Kimmel and others 1990). Arhonditsis and others (2007) found that along a freshwater-estuarine gradient, phytoplankton growth and community structure were strongly affected by river flow fluctuations. The direction of the effect was dependent upon location along the river to estuary gradient, with negative effects during high flow rates at river sites (indicating advective losses) to positive regulation closer to the estuary. Similarly, discharge rates in small lakes have been shown to be more important in predicting phytoplankton than nutrients, light and temperature (Dickman 1969).

In our study, RDA and short-term HRT was moderately correlated ($r = -0.60$); however, HRT was not identified by either model as a significant predictor of production or N_2 fixation. Although our measure of HRT may best describe aquatic conditions immediately preceding and during the sampling, it did not have predictive power on a spatial scale. Other studies concerned with broad spatial patterns of phytoplankton production have used average annual HRTs (Jones and Elliott 2007; Søballe and Kimmel 1987). The latter study found that both the strength and direction of correlations between HRT and drainage area (both log transformed) were fundamentally different for rivers ($r = 0.98$), impoundments ($r = -0.19$) or natural

lakes ($r = -0.34$), suggesting that in reservoirs, the relationship between drainage area and HRT would be expected to change from positive to negative along the longitudinal gradient from riverine to lacustrine zones. In our CART model for phytoplankton production, smaller RDAs (longer HRTs and reduced flushing) were associated with higher production rates. Although these results underline the interaction between hydrodynamics and location within the reservoir, the usefulness of RDA as a proxy for hydrologically driven processes needs further investigation.

Nitrogen Fixation Models

Prediction of nitrogen fixation in lakes and reservoirs is considerably more complex than predictions of phytoplankton production because nitrogen fixation depends both on succession of specific cyanobacteria, and factors that initiate the formation of heterocysts (Howarth and others 1988a). Two broad trends that have emerged are (1) rates of fixation are reasonably correlated with the biomass of N-fixing cyanobacteria (Wetzel 1983; Goldman and Horne 1983); and (2) rates tend to be higher in enriched lakes (Howarth and others 1988a). Additional factors that have been proposed, but whose roles are less clear, include both biogeochemical controls such as nutrients and bioavailable nitrogen, and physical controls such as turbulence, light intensity, and hydraulic residence time (Howarth and others 1988b).

In our study, the most important predictor of potential N_2 fixation in both linear and regression tree models was RDA. As previously discussed, sites with small RDAs would presumably experience more frequent periods of stagnation, which has been found to override nutrient levels as a controlling factor of N_2 fixing cyanobacteria (Paerl 1985). Theoretically, higher flushing associated with larger RDAs could result in negative effects such as higher turbidities and turbulence. Paerl (1985) found that turbulence was associated with reductions in nitrogen fixation because it hampered the development of cyanobacteria-bacterial aggregates. Patoine and others (2006) measured nitrogen fixation in six Canadian prairie lakes connected along a hydrologic gradient and found that N_2 fixation was significantly correlated to both landscape position and effective drainage area (defined as the region supplying water to a lake during years of median river flow). However, catchment:lake area ratios and mean flushing rates (y^{-1}) were not statistically significant predictors.

Arhonditsis and others (2007) found that along a freshwater-estuarine gradient, cyanobacteria growth was higher in areas with longer HRTs which also had relaxed phosphorus limitation. Thus, longer retention times may favor N_2 fixing cyanobacteria because they are slower growing and thus may be at a disadvantage in faster flushing systems (Maberly and others 2002).

The role of bioavailable nitrogen such as dissolved inorganic nitrogen and some organic nitrogen forms has been documented in seasonal initiation and suppression of nitrogen fixation (Scott and others 2008; Zevenboom and Mur 1980). In our study, neither multiple linear regression nor the pruned regression tree analyses detected a significant relationship between bioavailable nitrogen (that is, nitrate) and N_2 fixation potentials. This is not entirely surprising, however, as the roles of bioavailable nitrogen, as well as temperature, are normally revealed in the seasonal onset of N_2 fixation (for example, Scott and others 2008). However, we did detect higher nitrate levels at both Lewisville and Texana reservoirs which may also have contributed to low fixation rates at these sites. In fact, Lewisville and Texana had mean nitrate levels (31 and 34 $\mu g\ l^{-1}$, respectively) at least six times higher than the remaining reservoirs ($2.7–5.1\ \mu g\ l^{-1}$) which may have contributed to suppression of nitrogen fixation. This is slightly below the threshold levels of below 50–100 $\mu g\ N\ l^{-1}$ dissolved inorganic nitrogen proposed by Horne and Commins (1987), but consistent with threshold of 25 $\mu g\ N\ l^{-1}$ in a nearby central Texas reservoir identified by Scott and others (2008). Conversely, other lake studies have documented nitrogen fixation in the presence of substantial quantities of ammonium, and found that strong relationships between dissolved inorganic nitrogen and rates of N_2 fixation have been lacking (Paerl and others 1981).

Many studies have reported a positive relationship between phosphorus concentrations and N_2 fixation (Vanderhoef and others 1974; Sterner 1994; Arhonditsis and others 2007). Smith (1983) examined N_2 fixation data from 17 freshwater lakes and reservoirs spanning a large geographic area and found that there was a highly significant ($P < 0.001$) unimodal relationship between annual rates of nitrogen fixation and mean growing season TP, with peak rates of nitrogen fixation occurring at intermediate concentrations of TP. Our N_2 fixation potentials also exhibited a unimodal relationship with TP (data not shown). For example, Lewisville-B, which had the study's highest TP (mean 719 $\mu g\ l^{-1}$), and Lake Texana, which had

TPs greater than 100 $\mu g\ l^{-1}$, exhibited minimal N_2 fixation potentials. Our CART model found that N_2 fixation potential was positively related to TP only at sites with smaller RDA, suggesting that nutrient controls are spatially distributed and subordinate to physical factors such as hydrological flushing. This is consistent with observations along a riverine–estuarine where cyanobacteria growth was higher in downstream areas (estuarine and transitional zone) due to longer residence times, transport from up-river, lower dissolved inorganic nitrogen, and relaxed phosphorus limitation (Arhonditsis and others 2007).

The ratio of N:P has been considered an important predictor of dominance by N_2-fixing cyanobacteria in lakes. The N:P that facilitates dominance by N_2-fixers has been reported to vary from values below the Redfield Ratio of 15.5 (molar) to values of 20–40 (44–88 molar) in shallow Estonian lakes (Nõges and others 2008). In a central Texas reservoir, Scott and others (2008) used CART analyses to identify a threshold N:P (from watershed loads) of 32, below which N_2 fixation rates were higher, but only at sites located in one arm of the reservoir. The effect was minimal at sites located in a different arm or near the dam. Furthermore, the partial r^2 was small (0.16) compared to the variability that was explained by temperature (0.31). The N:P (molar TN:TP) in our study ranged from 4 to 340. Neither correlation nor linear regression identified N:P as the optimal predictor for where N_2 fixation rates would be high. However, N:P ratio (threshold = 28.5) was selected by CART as an alternative predictor for TP, which is consistent with the strong correlation between Log TP and Log N:P in our data ($r = -0.888$). N:P was also selected at a lower node (threshold = 28.7), but this split was pruned according to the one standard error rule as well as a low partial r^2 of 0.05. The N:P split occurred on the branch with RDAs below 45 and TP values above 40 $\mu g\ l^{-1}$. Additionally, N_2 fixation rates were only slightly higher at sites with N:P below 28.7.

Land use played a minor role in predicting phytoplankton production potential and no significant role in predicting N_2 fixation potential. However, it has been well established that increased development in watersheds is correlated with increased TP (Weibel 1969; Carpenter and others 1998b) and decreased N:P of hydraulic loads (Downing and McCauley 1992). In this study, proportion of undeveloped land and TP were related; however, TP levels were highly variable at sites with the largest proportions of undeveloped land (data not shown, $r^2 = 0.26$). This may be a function of insufficient replication, because our 85 sites were

actually represented by only 24 land-use distributions (Figure 2). Other factors such as effluent releases, proximity to degradation (King and others 2005a), and watershed differences in slope, water yield, and soils were not accounted for in our land use analyses. Thus, our results do not rule out potential effects associated with land use.

Another important limitation of our study was the lack of temporal replication. Nitrogen fixation as well as bioavailable nutrient levels are known to vary considerably even during summer months and it is possible that we failed to sample during periods of N_2 fixation in water bodies where N_2 fixation occurs. In addition, our study occurred during relatively low flow conditions and our results cannot address interannual variability associated with meterological processes.

Despite these limitations, our analyses indicate that regression tree modeling can be a valuable tool for predicting algal production and nitrogen fixation potentials along riverine to lacustrine gradients. We also found substantial differences between regression tree models and those developed by traditional linear regression. For example, our regression tree models revealed threshold values for predictor variables, which can be interpreted as decision nodes for predicting response values. Thus, CART may be more useful in evaluating where in the study area water quality problems are likely to occur. In contrast, multiple linear regression uses a single equation to predict responses regardless of location. In addition, multiple linear regression must fit all predictors simultaneously, and thus did not reveal the hierarchical structure among predictors in our data. Tree analyses work best when predictor–response relationships are nonlinear or heteroscedastic, properties that are inherent in many ecological data (Qian and others 2003). This was evidenced in the selection of depth by regression tree analyses as the most important predictor for phytoplankton production, whereas that predictor was overlooked by multiple linear regression. However, for strong linear relationships, threshold values will have higher uncertainty in tree type models (King and Richardson 2003). This was reflected in the high uncertainty estimates for TP in both regression tree models, which had a more linear relationship to the response variables in the tree models.

We also found that physical factors exhibited hierarchical control over nutrient effects, which is generally consistent with recent spatial scale studies of phytoplankton production (Søballe and Kimmel 1987; Cuevas and others 2006; Arhonditsis and others 2007) and N_2 fixation processes (Patoine and others 2006; Scott and others 2008; Arhonditsis and others 2007). Furthermore, we concluded that these relationships changed along environmental gradients associated with reservoir zones. Although the usefulness of large-scale predictors such as RDA require further examination, we found that this proxy for hydrologic processes was well correlated to N_2 fixation potential, and moderately correlated to phytoplankton production.

ACKNOWLEDGMENTS

REFERENCES

Adamus PR, Stockwell CT. 1983. A method for wetland function assessment, volume I: critical view and evaluation concepts. United States Federal Highway Administration, FHWA-IP-82-23.

Arbuckle KE, Downing JA. 2001. The influence of watershed land use on lake N:P in a predominantly agricultural landscape. Limnol Oceanogr 46:970–5.

Arhonditsis GB, Stow CA, Paerl HW, Valdes-Weaver LM, Steinberg LJ, Reckhow KH. 2007. Delineation of the role of nutrient dynamics and hydrologic forcing on phytoplankton patterns along a freshwater–marine continuum. Ecol Model 208:230–46.

Berman T. 2001. The role of DON and the effect of N:P ratios on occurrence of cyanobacterial blooms: implications from the outgrowth of *Aphanizomenon* in Lake Kinneret. Limnol Oceanogr 46:443–7.

Breiman L, Friedman JH, Olshen RA, Stone CJ. 1984. Classification and regression trees. Belmont (CA): Wadsworth International Group. 358 pp.

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, Smith VH. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol Appl 8:559–68.

Carpenter SR, Cole JJ, Hodgson JR, Kitchell MF, Pace ML, Bade D, Cottingham KL, Essington TE, Houser JN, Schindler DE. 2001. Trophic cascades, nutrients, and lake productivity: whole-lake experiments. Ecol Monogr 71:163–86.

Carpenter SR, Cole JJ, Kitchell JF, Pace ML. 1998. Impact of dissolved organic carbon, phosphorus and grazing on phytoplankton biomass and production in experimental lakes. Limnol Oceanogr 43:73–80.

Cuevas JG, Soto D, Arismendi I, Pino M, Lara A, Oyarzún C. 2006. Relating land cover to stream properties in southern Chilean watersheds: trade-off between geographic scale, sample size, and explicative power. Biogeochemistry 81: 313–29.

De'Ath G, Fabricius KE. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178–92.

Dickman M. 1969. Some effects of lake renewal on phytoplankton productivity and species composition. Limnol Oceanogr 14:660–6.

Downing JA, McCauley E. 1992. The nitrogen:phosphorus relationship in lakes. Limnol Oceanogr 37:936–45.

Fee EJ. 1973. A numerical model for determining integral primary production and its application to Lake Michigan. J Fish Res Board Can 30:1447–68.

Flett RJ, Hamilton RD, Campbell NER. 1975. Aquatic acetylene-reduction techniques: solutions to several problems. Can J Microbiol 22:43–51.

Gilbert MC, Whited PM, Clairain EJ, Jr, Smith RD. 2006. A regional guidebook for applying the hydrogeomorphic approach to assessing wetland functions of prairie potholes. ERDC/EL TR-06-5.

Goldman CR, Horne AJ. 1983. Limnology. New York, USA: McGraw-Hill.

Gordon JA, Behel RM. 1985. Suspended sediment characteristics of Lake Cumberland, Kentucky. In: Proceedings North American Lake Management Society, 1984. p 259–64.

Grobbelaar JU. 1989. The contribution of phytoplankton productivity in turbid freshwaters to their trophic status. Hydrobiologia 173:127–33.

Harmel RD, Richardson CW, King KW, Allen PM. 2006. Runoff and soil loss relationships for the Texas Blackland Prairies ecoregion. J Hydrol 331:471–83.

Horne AJ, Commins ML. 1987. Macronutrient controls on nitrogen fixation in planktonic cyanobacterial populations. N Z J Marine Freshw Res 21:413–23.

Horppila J, Liljendahl-Nurminen A. 2005. Clay–turbidity interactions may not cascade—a reminder for lake managers. Restorat Ecol 13:242–6.

Howarth RW, Marino R, Lane J, Cole JJ. 1988. Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 1. Rates and importance. Limnol Oceanogr 33:669–87.

Howarth RW, Marino R, Cole JJ. 1988. Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 2. Biogeochemical controls. Limnol Oceanogr 33:688–701.

Huppert A, Blasius B, Stone L. 2002. A model of phytoplankton blooms. Am Nat 159:156–71.

Jones ID, Elliott JA. 2007. Modeling the effects of changing retention time on abundance and composition of phytoplankton species in a small lake. Freshw Biol 52:988–97.

Kennedy RH, Walker WW. 1990. Reservoir nutrient dynamics. In: Thornton K, Kimmel B, Payne F, Eds. Reservoir limnology: ecological perspectives. New York: Wiley & Sons. p 109–32.

Kimmel BL, Lind OT, Paulson LJ. (1990). Reservoir primary production. In: Thornton K, Kimmel B, Payne F, Eds. Reservoir limnology: ecological perspectives. New York: Wiley & Sons. p 133–94.

King RS, Deluca WV, Whigham DF, Marra PP. 2007. Threshold effects of coastal urbanization on *Phragmites australis* (common reed) abundance and foliar N in Chesapeake Bay. Estuaries Coasts 30:469–81.

King RS, Hines AH, Craige FD, Grap S. 2005. Regional, watershed and local correlates of blue crab and bivalve abundances in subestuaries of Chesapeake Bay, USA. J Exp Marine Biol Ecol 312:101–16.

King RS, Baker ME, Whigham DF, Weller DE, Jordan TE, Kazyak PF, Hurd MK. 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecol Appl 15:137–53.

Levine SN, Lewis WM Jr. 1987. A numerical model of nitrogen fixation and its application to Lake Valencia, Venezuela. Freshw Biol 17:265–74.

Lewis RJ. 2000. An Introduction to classification and regression tree (CART) analysis. In: Presented at the 2000 annual meeting of the Society for Academic Emergency Medicine in San Francisco, California.

Lind OT, Doyle R, Vodopich DS, Trotter BG, Limón JG, Dávalos-Lind L. 1992. Clay turbidity: regulation of phytoplankton production in a large, nutrient-rich tropical lake. Limnol Oceanogr 37:549–65.

Maberly SC, King L, Dent MM, Jones RI, Gibson CE. 2002. Nutrient limitation of phytoplankton and periphyton growth in upland lakes. Freshw Biol 47:2136–52.

McQueen DJ, Lean DRS. 1987. Influence of water temperature and nitrogen to phosphorus ratios on the dominance of blue-green algae in Lake St. George, Ontario. Can J Fish Aquat Sci 44:598–604.

Nõges T, Laugaste R, Nõges P, Tõnno I. 2008. Critical N:P ratio for cyanobacteria and N_2-fixing species in the large shallow temperate lakes Peipsi and Võrtsjärv, North-East Europe. Hydrobiologia 599:77–86.

Osidele OO, Beck MB. 2004. Food web modelling for investigating ecosystem behaviour in large reservoirs of the southeastern United States: lessons from Lake Lanier, Georgia. Ecol Model 173:129–58.

Paerl HW. 1985. Microzone formation: its role in the enhancement of aquatic N_2 fixation. Limnol Oceanogr 30:1246–52.

Paerl HW, Webb KL, Baker J, Wiebe WJ. 1981. Nitrogen fixation in natural waters. In: Broughton, WJ, Ed. Nitrogen fixation, volume 1, Ecology. Oxford: Clarendon Press. p 193–240.

Parr LB, Mason CF. 2004. Causes of low oxygen in a lowland, regulated eutrophic river in Eastern England. Sci Total Environ 321:273–86.

Patoine A, Graham MD, Leavitt PR. 2006. Spatial variation of nitrogen fixation in lakes of the northern Great Plains. Limnol Oceanogr 51:1665–77.

Pickett JR, Harvey RM. 1988. Water quality gradients in the Santee-Cooper Lakes, South Carolina. Lake Reserv Manag 4:11–20.

Qian SS, King RS, Richardson CJ. 2003. Two statistical methods for the detection of environmental thresholds. Ecol Model 166:87–97.

Rejwan C, Collins NC, Brunner LJ, Shuter BJ, Ridgway MS. 1999. Tree regression analyses on the nesting habitat of smallmouth bass. Ecology 80:341–8.

Schindler DW. 1978. Factors regulating phytoplankton production and standing crop in the world's freshwaters. Limnol Oceanogr 23:148–486.

Scott JT, Doyle RD, Prochnow S, White JD. 2008. Are watershed and lacustrine controls on planktonic N^2 fixation hierarchically structured? Ecol Appl 18:805–19.

Smayda TJ. 1970. The suspension and sinking of phytoplankton in the sea. Oceanogr Marine Biol Annl Rev 8:353–414.

Smith VH. 1983. Low nitrogen to phosphorus ratios favor dominance by blue–green algae in lake phytoplankton. Science 221:669–71.

Smith VH. 1990. Nitrogen, phosphorus, and nitrogen fixation in lacustrine and estuarine ecosystems. Limnol Oceanogr 35:1852–59.

Smith RK, Baker KS, Holm-Hansen O, Olson R. 1980. Photoinhibition of photosynthesis in natural waters. J Photochem Photobiol 31:585–92.

Smith VH, Shapiro J. 1981. Chlorophyll–phosphorus relations in individual lakes: their importance to lake restoration strategies. Environ Sci Technol 15:444–51.

Søballe DM, Kimmel BL. 1987. A large-scale comparison of factors influencing phytoplankton abundance in rivers, lakes, and impoundments. Ecology 68:1943–54.

Sterner RW. 1989. Resource competition during seasonal succession toward dominance by cyanobacteria. Ecology 10: 229–45.

Sterner RW. 1994. Seasonal and spatial patterns in macro- and micronutrient limitation in Joe Pool Lake, Texas. Limnol Oceanogr 39:535–50.

Thomas DB, Schallenberg M. 2007. Benthic shear stress gradient defines three mutually exclusive modes of non-biological internal nutrient loading in shallow lakes. Hydrobiologia 610:1–11.

Thornton KW. 1990. Perspectives on reservoir limnology. In: Thornton K, Kimmel B, Payne F, Eds. Reservoir limnology: ecological perspectives. New York: Wiley and Sons. p 1–14.

Thornton KW, Kennedy RH, Carroll JH, Walker WW, Gunkel RC, Ashby S. 1981. Reservoir sedimentation and water quality—an heuristic model. In: Stefan HG, Ed. Proceedings of the symposium on surface water impoundments. New York (NY): American Society of Civil Engineers.

Thornton KW, Kimmel BL, Payne FE, Eds. 1990. Reservoir limnology: ecological perspectives. Wiley & Sons, Inc. 245 pp.

Urban DL. 2002. Classification and regression trees. In: McCune B, Grace JB, Eds. Analysis of ecological communities. Oregon: MjM Software Design. p 222–32.

Vanderhoef LN, Huang C, Musil R, Williams J. 1974. Nitrogen fixation (acetylene reduction) by phytoplankton in Green Bay, Lake Michigan, in relation to nutrient concentrations. Limnol Oceanogr 19:119–25.

Vanni MJ, Andrew JS, Renwick WH, Gonzales MJ, Noble SJ. 2006. Nutrient and light limitation of reservoir phytoplankton in relation to storm-mediated pulses in stream discharge. Arch. Hydrobiol. WL37:1–25.

Venables WN, Ripley BD. 1994. Modern applied statistics with S-plus. New York: Springer, p 501.

Vollenveider RA, Kerekes JJ. 1980. Background and summary results of the OECD Cooperative Program on Monitoring of Inland Waters (Eutrophication Control). p 290.

Wetzel RG. 1983. Limnology. 2nd edn. Orlando, FL: Saunders College Publishing, p 756.

Weibel SR. 1969. Urban drainage as a factor in eutrophication. Eutrophication: causes, consequences, correctives. Washington, DC: National Academy of Sciences, p 383–403.

Zevenboom W, Mur LR. 1980. N_2-fixing cyanobacteria: why they do not become dominant in Dutch, hypertrophic lakes. In: Barica J, Mur LR, Eds. Hypertrophic ecosystems. The Hague, Netherlands: Dr. W. bv Junk Publishing. p 123–30.