

A Bayesian network model for estimating stoichiometric ratios of lake seston components

Lester L. Yuan ^a and John R. Jones^b

^aOffice of Water, US Environmental Protection Agency, Washington, DC, USA; ^bSchool of Natural Resources, University of Missouri – Columbia, Columbia, MO, USA

ABSTRACT

The elemental composition of seston provides insights into how lake food webs function and how nutrients cycle through the environment. Here, we describe a Bayesian network model that simultaneously estimates relationships between dissolved and particulate nutrients, suspended volatile and nonvolatile sediments, and algal chlorophyll. The model provides direct estimates of the phosphorus (P) and nitrogen (N) content of phytoplankton, suspended non-living organic matter, and suspended inorganic sediment. We applied this model to data collected from reservoirs in Missouri, USA, to test the validity of our assumed relationships. The results indicate that, on average among all samples, the ratio of N and P (N:P) in phytoplankton and non-living organic matter in these reservoirs were similar, although under nutrient replete conditions, N:P in phytoplankton decreased. P content of inorganic sediment was lower than in phytoplankton and non-living organic matter. The analysis also provided a means of tracking changes in the composition of whole seston over time. In addition to informing questions regarding seston stoichiometry, this modeling approach may inform efforts to manage lake eutrophication because it can improve traditional models of relationships between nutrients and chlorophyll in lakes.

ARTICLE HISTORY

Received 24 May 2018
Revised 28 January 2019
Accepted 5 February 2019

KEYWORDS

Bayesian model; nitrogen; phosphorus; seston; stoichiometry

Introduction

Since Redfield's seminal work (Redfield 1958), aquatic ecologists have measured the elemental composition of seston and sought to explain its relationship to processes such as nutrient cycling (Elser and Urabe 1999), nutrient limitation for both phytoplankton (Klausmeier et al. 2004b, Ptacnik et al. 2010, Hillebrand et al. 2013) and herbivores (Sterner and Hessen 1994), and food web stability (Perhar and Arhonditsis 2009). In general, elemental composition of seston is thought to reflect the basic processes that cycle carbon (C), phosphorus (P), and nitrogen (N) in the environment (Sterner et al. 2008). Because human activities have significantly altered nutrient concentrations in aquatic ecosystems, a better understanding of the causes of changes to the elemental composition of seston is critical for managing and mitigating the effects of excess nutrients in the environment (Dodds 2006, Glibert et al. 2011).

Partitioning seston into its components, which include living phytoplankton, non-living organic matter, and inorganic matter, is difficult using data routinely collected from field studies. Many studies rely on a series of filters to separate particles of different size ranges and then assume each size range is composed primarily of

unique aspects of seston, such as phytoplankton (Elser and Hassett 1994) or bacteria (Cotner et al. 2010). In some studies, post hoc tests are used to assess accuracy of these initial assumptions (Hassett et al. 1997). Particle size is, however, imperfectly associated with seston type. Consequently, analyses of elemental composition of bulk seston, even in restricted size ranges, provide an incomplete picture of nutrient allocation among various components of the seston. For example, the range of particle sizes associated with inorganic matter overlaps with particle sizes associated with organic matter (Walling and Moorehead 1989).

Various laboratory methods for separating different components of seston are currently used. For example, centrifuging samples combined with colloidal silica separates living from non-living organic matter (Hamilton et al. 2005). Alternatively, measurements of inorganic particle density via electron microscopy and x-ray analyses can partition total P measurements into organic and inorganic components (Effler et al. 2014). These methods are promising but not routinely applied.

Other workers have proposed statistical approaches for separating the components of seston (Hessen et al. 2003), and we expand on this approach. More specifically, we asked whether the elemental composition of

living phytoplankton, non-living organic matter, and inorganic matter can be inferred by invoking basic assumptions and analyzing measurements of total phosphorus (TP), total nitrogen (TN), suspended sediment, and chlorophyll (Chl). We used a Bayesian network model for this analysis that provides specific advantages (described later) over more commonly applied frequentist statistical methods. We then demonstrated that partitioning seston components in this way yields new insights into mechanisms driving changes in seston stoichiometry and potentially improve our ability to predict Chl from nutrient concentrations.

Methods

Data

We illustrated the models with data collected in 15 reservoirs constructed in the Glacial Till Plains region of northern Missouri, USA, between 1950 and 1992. The range in size (10–408 ha), mean depth (1.6–5.7 m), and flushing rate (0.1–3.4 times per year) represents regional reservoir conditions (Jones et al. 2008). Integrated photic zone samples were collected weekly near the dam of each reservoir between May and August 2004. Total suspended solids (TSS) were determined by filtering a known volume of lake water through Whatman934-AH filters (nominal filter size: 1.5 μm ; Whatman, Maidstone, UK) that were pre-rinsed, dried, ashed, and tared. Nonvolatile suspended sediment (NVSS) was determined by weight after ashing TSS samples, and volatile suspended sediment (VSS) was determined by difference (TSS – NVSS). Samples were analyzed for Chl (uncorrected for degradation products), TP, dissolved P, TN, dissolved inorganic N (DIN), and dissolved N. Chl was measured from material retained on a 1 μm Gelman AE filter (Pall Corporation, NY, USA). Dissolved nutrient concentrations were estimated in filtrate through the Whatman934-AH filters. A total of 241 samples were available for analysis.

Data collected by the US Environmental Protection Agency (EPA) National Lakes Assessment (NLA) in summers (May–Sep) 2007 and 2012 were used to qualitatively compare with trends estimated from Missouri reservoirs. The NLA consisted of a random sample of ~ 1800 lakes from the continental United States. During each visit to a selected lake, a sampling location was established in open water at the deepest point of each lake (up to a maximum depth of 50 m) or in the mid-point of reservoirs. A water sample was collected using a vertical, depth-integrated methodology that collected water from the photic zone of the lake (to a maximum depth of 2 m). Multiple sample draws were combined

in a rinsed, 4 L cubitainer. When full, the cubitainer was gently inverted to mix the water, and an aliquot was taken as the water chemistry sample. This subsample was placed on ice and shipped overnight to the Willamette Research Station (Corvallis, OR, USA). TN and TP were measured in the laboratory from the open-water sample at pre-specified levels of precision and accuracy (US EPA 2012). Typical laboratory methods included persulfate digestion with colorimetric analysis for TN and TP. To measure Chl concentration, 250 mL of lake water was pumped through a glass fiber filter in the field and quantified in the lab to pre-specified levels of precision and accuracy.

Mathematical model

For TP, we assumed that 3 seston components in addition to dissolved P contribute to TP: P sorbed onto inorganic (or NVSS), P associated with living phytoplankton, and P associated with non-living organic suspended sediment (VSS_{np} ; Fig. 1).

Direct measurements of VSS_{np} were not available, but we estimated its concentration by assuming that the concentration of all volatile suspended sediments is composed of 2 main contributors: phytoplankton biomass and non-living organic matter not associated with living phytoplankton (i.e., VSS_{np}). Hence, we modeled observed concentrations of VSS as follows:

$$VSS = kChl + VSS_{np}, \quad (1)$$

where VSS concentration is expressed as the sum of a phytoplankton component, modeled as the product of a coefficient (k) multiplied by Chl concentration, and a non-phytoplankton component, VSS_{np} . The distribution

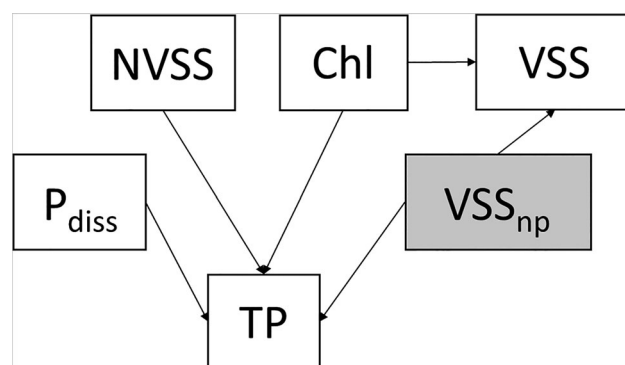


Figure 1. Schematic representation of the model for the Chl–TP relationship. Observed TP originates from phytoplankton, which is associated with chlorophyll (Chl), dissolved P (P_{diss}), sediment-bound P (NVSS), and P associated with non-living organic matter (VSS_{np}). The shaded box for VSS_{np} indicates that direct measurements of this parameter were not available, and its value is inferred statistically.

of k is assumed to be log-normal to constrain it to positive values. To fit this relationship to available data, we log-transformed both sides of the equation to address skewness in the observed distributions of the different variables, and we introduced $\varepsilon_{1,i}$, a term that allows residual variability in observed $\log(\text{VSS})$:

$$\log(\text{VSS}_i) = \log(k\text{Chl}_i + \text{VSS}_{\text{np},i}) + \varepsilon_{1,i}, \quad (2)$$

where the subscript i refers to individual samples in the dataset, and the error term $\varepsilon_{1,i}$ is a random variable drawn from a normal distribution with a mean value of zero and a standard deviation of σ_1 . Exploratory analysis suggested that mean values for VSS_{np} varied among lakes, so values of $\text{VSS}_{\text{np},i}$ were assumed to be drawn from a log-normal distributions characterized by a mean, $\mu_{\text{VSS-np}}$, unique to each lake and a standard deviation, $\sigma_{\text{VSS-np}}$:

$$\log(\text{VSS}_{\text{np},i}) = N(\mu_{\text{VSS-np},j[i]}, \sigma_{\text{VSS-np}}), \quad (3)$$

where the subscript $j[i]$ refers to different lakes that correspond with each sample, i . Mean values for each lake ($\mu_{\text{VSS-np},j}$) were assumed to be drawn from a single common normal distribution, characterized by an overall mean value and standard deviation.

Observed TP is modeled as the sum of directly observed dissolved P, P incorporated in phytoplankton (assumed to be directly proportional to Chl), P associated with non-living organic matter (VSS_{np}), and NVSS:

$$E[\text{TP}_i] = P_{\text{diss},i} + d_1\text{Chl}_i + d_2\text{VSS}_{\text{np},i} + d_3\text{NVSS}_i. \quad (4)$$

The 3 coefficients, d_1 , d_2 , and d_3 , quantify the mass of P in phytoplankton per unit of Chl and the P content of VSS_{np} and NVSS, respectively. Like k , these coefficients are assumed to be log-normally distributed to restrict them to positive values. The subscript i again refers to individual samples, and TP_i is enclosed by $E[\cdot]$ to indicate that the equation applies to the expected value of TP for sample i . In contrast to the model for VSS, we did not log-transform the equation. Instead, the skewed distribution of observed values for TP_i was taken into account by modeling the residual distribution of these observed values as a gamma distribution with a shape parameter of α_1 and a rate parameter of $\alpha_1/E[\text{TP}_i]$. The mean value of this distribution is $E[\text{TP}_i]$ and the variance is $E[\text{TP}_i]^2/\alpha_1$, so the residual variance is proportional to the square of the expected value. The parameter α_1 adjusts the rate at which residual variance increases with the mean expected value and is fit to the observed data.

We specified an identical model for TN, but initial model fits indicated that the contribution of NVSS to TN was negligible, and so we excluded this term in the

final model:

$$E[\text{TN}_i] = N_{\text{diss},i} + f_1\text{Chl}_i + f_2\text{VSS}_{\text{np},i}. \quad (5)$$

Here again, the parameters f_1 and f_2 are specified as log-normally distributed to constrain them to positive values, and observed values of TN are assumed to be drawn from a gamma distribution with a shape parameter of α_2 .

To further explore the potential for this type of analysis to provide insights regarding the mechanisms that determine stoichiometric ratios, we identified Missouri samples with DIN $>50 \mu\text{g/L}$. In these samples, sufficient N was likely available for phytoplankton growth, and therefore we hypothesized that stoichiometric ratios associated with phytoplankton would differ in these samples compared to samples in which DIN was near zero (Klausmeier et al. 2004a). We refit the models but estimated different values for the coefficients that quantified N- and P-content of phytoplankton and VSS_{np} . More specifically, we expressed our model equations as follows:

$$E[\text{TP}_i] = P_{\text{diss},i} + d_{1,k}\text{Chl}_i + d_{2,k}\text{VSS}_{\text{np},i} + d_3\text{NVSS}_i, \quad (6)$$

$$E[\text{TN}_i] = N_{\text{diss},i} + f_{1,k}\text{Chl}_i + f_{2,k}\text{VSS}_{\text{np},i}, \quad (7)$$

where the index k indexes samples with low ($k=1$) and high ($k=2$) concentrations of DIN.

Prior to analysis, all measurements were scaled by dividing by their maximum observed value to improve the efficiency of the Monte-Carlo algorithm used to fit the equation. We then specified weakly informative priors for all the parameters. As discussed earlier, all model coefficients, k , f_1 , f_2 , d_1 , d_2 , and d_3 , were assumed to be log-normally distributed to restrict them to positive values. Prior distributions for these parameters were specified as log-normal distributions with mean values of 0 and standard deviations of 8. The value of 8 was chosen to be several times larger than the expected standard deviation of any of the coefficients so that the function of the prior distribution was mainly to steer randomly drawn samples away from extreme values. Similarly, the prior distribution of σ_1 was specified as a half-Cauchy distribution with a scale of 2. Initial estimates of the parameters α_1 and α_2 were large, and so to improve the efficiency of the Monte-Carlo sampling we scaled α_1 and α_2 by a factor of 100 and assigned them weakly informative half-Cauchy priors with a scale of 2. The prior distribution for the mean values of VSS_{np} (i.e., $\mu_{\text{VSS-np}}$) was a normal distribution with a mean value of 0 and a standard deviation of 8. All models were then fit simultaneously using the Bayesian modeling software, RStan (Stan Development Team 2016), and the statistical modeling software R 3.4.2 (R Core Team 2017).

Because the relationships estimated in the current models quantify inherent mean stoichiometric relationships regarding P and N content in phytoplankton cells, these relationships may be similar among different datasets. To qualitatively test this hypothesis, we compared base relationships estimated in Missouri to TP, TN, and Chl data collected from lakes across the United States (US EPA 2010, 2017).

Results

The TP and TN models provided accurate predictions of observed TP and TN (Fig. 2). Because residual variance increased with the mean predicted values of TP and TN, root mean square prediction error (RMSE) exhibited similar trends, ranging from 4 to ~ 20 $\mu\text{g/L}$ for the TP model, and from 50 to 200 $\mu\text{g/L}$ for TN, mirroring the predicted increase in residual standard deviation for the gamma distribution (solid lines in lower panels of Fig. 2). We also computed 90% posterior prediction intervals for each sample and confirmed that $<10\%$ of observations of TP and TN deviated from these intervals.

Coefficient values for the base model indicated that on average, NVSS was 0.0999% P by weight, whereas N content was negligible (Table 1). Mean P and N content in non-phytoplankton VSS was 0.67% and 6.9%, respectively, which yielded a molar ratio of N:P in VSS_{np} of

Table 1. Estimates for model coefficients for the Bayesian network model. Equation numbers for each model correspond to those in the text.

Model	Coefficient	Mean value (90% confidence limits)
VSS (Eq. 2)	k	0.092 (0.084, 0.10)
TP (Eq. 4)	d_1	0.61 (0.54, 0.68)
	d_2	6.7 (5.9, 7.7)
	d_3	0.99 (0.70, 1.28)
TN (Eq. 5)	f_1	5.4 (4.7, 6.1)
	f_2	69 (61, 78)

23:1; 90% credible intervals on this ratio extended from 19:1 to 27:1.

Coefficients for Chl in the TN and TP models quantified the ratio between the mass of N and P in phytoplankton and the mass of Chl. So, we estimated the mean mass ratio between N and P within phytoplankton cells as $5.4/0.61 = 8.8$. This mass ratio converts to a molar ratio of 19 in phytoplankton, with 90% credible intervals extending from 17 to 23.

The coefficient k in the model for VSS measures the proportion of the observed VSS mass attributable to phytoplankton (as quantified by Chl). If we assume that VSS concentration can be converted to an estimate of C by multiplying by a factor of 0.45 (Findlay et al. 1991, Pribyl 2010), we estimate the molar C:P ratio in phytoplankton as 170:1, with 90% credible intervals of 160:1 and 200:1.

Distributions of TP and TN measurements are highly skewed, and so these variables are frequently

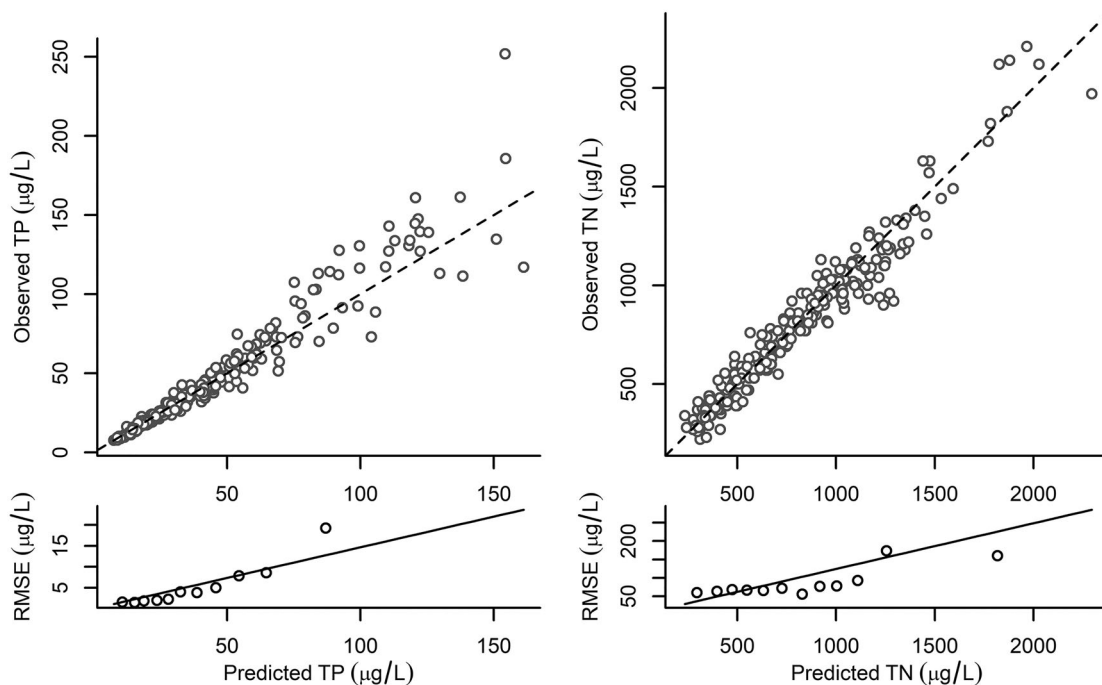


Figure 2. Predicted TP and TN versus observed TP and TN (upper panels). Root mean square error (RMSE) as a function of predicted values (lower 2 panels). Dashed lines (upper panels) show 1:1 relationships. Solid line in lower 2 panels shows the increase in residual standard deviation expected from gamma distribution. Open circles in lower panels show estimated RMSE computed from ~ 20 samples in a bin around the indicated predicted TP or TN.

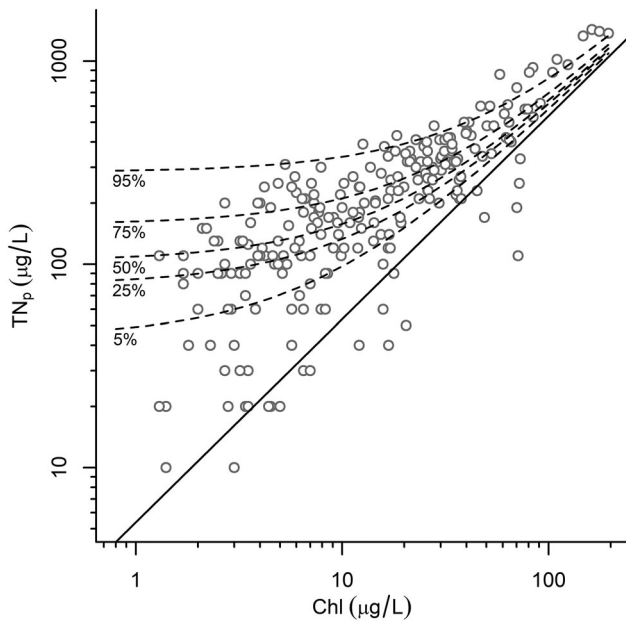


Figure 3. Predicted relationships between Chl and particulate TN for different levels of VSS_{np} . Solid line: relationships between Chl and TN_p when VSS_{np} is negligible. Dashed lines: relationship between Chl and TN_p for different quantiles of VSS_{np} (as indicated).

log-transformed prior to analysis, as are measurements of different predictor variables (e.g., Chl). In the present model, the relationships are expressed in terms of the variables in their original units, and therefore the log-transformation would apply to the entire sum of model components. For example, a log-transform applied to the model for TN would be:

$$\log(E[TN_i]) = \log(N_{diss,i} + f_1 Chl_i + f_2 VSS_{np,i}). \quad (8)$$

Hence, modeled relationships between a predictor variable (e.g., Chl) and $E[TN]$ are nonlinear when plotted in log-log space. Examples of this nonlinear relationship

are illustrated when Chl is plotted against particulate TN (defined as the difference between TN and dissolved N; Fig. 3). A limiting relationship between Chl and TN can be computed by considering the case in which VSS_{np} is negligible, in which case we assume that the term $f_2 VSS_{np}$ is zero and write the log-transformed model equation as follows:

$$\log(E[TN_p]) = \log(f_1) + \log(Chl), \quad (9)$$

where TN_p is particulate TN. So, in samples in which VSS_{np} is small relative to Chl, the relationship between $\log(Chl)$ and $\log(TN)$ is a straight line with a slope of 1 and an intercept value of $\log(f_1)$ (Fig. 3, solid line).

As the contribution of N associated with VSS_{np} increases, the relationship between Chl and TN_p deviates from the limiting relationship. For example, if we fix VSS_{np} at the fifth percentile of its observed distribution of values, the resulting relationship between Chl and TN_p approaches the limiting relationship quickly as Chl increases, but at low concentrations of Chl, the effect of VSS_{np} is strong enough to reduce the slope between Chl and TN_p to nearly zero (Fig. 3, “5%” dashed line). For relationships computed using higher concentrations of VSS_{np} , the curves begin to deviate more substantially from the limiting relationship but always approach the linear limit at high concentrations of Chl. Overall, the distribution of curves matches the distribution of the observed samples, lending support to the appropriateness of the model formulation. The same limiting relationships exist for the relationships between Chl and VSS and between Chl and TP (Fig. 4).

Seston contributions to TP and TN varied with time and by lake. In the example of an oligo/mesotrophic lake (left panel), dissolved N accounted for ~80% of observed TN throughout the summer sampling season

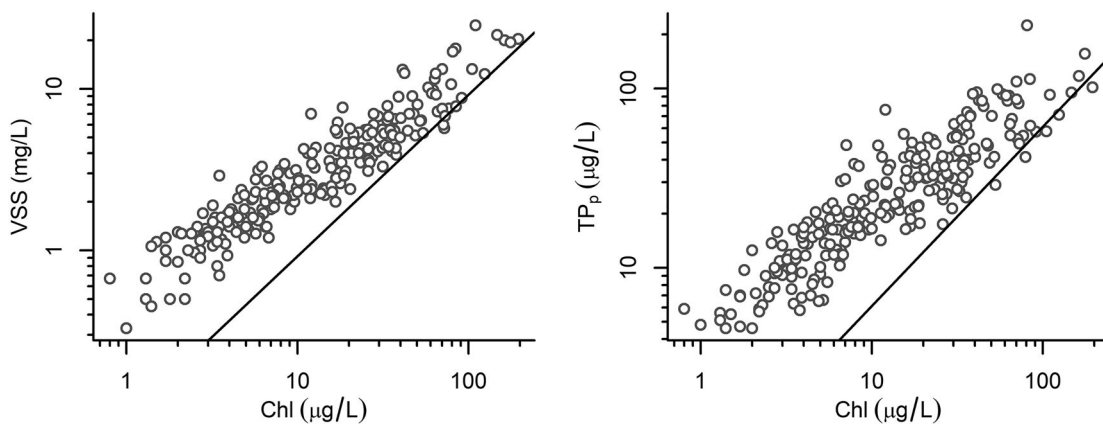


Figure 4. Relationships between Chl and VSS and TP_p . Solid lines show estimated limiting relationship between Chl and VSS (left panel) and between Chl and particulate TP (right panel) in the absence of contributions from any other factors.

(Fig. 5). Non-living organic seston accounted for the majority of the remaining TN, with only a small increase in phytoplankton-associated N in weeks 24 and 25. Dissolved P accounted for ~50% of observed TP throughout the summer in the same lake, and the proportional contribution of phytoplankton P was nearly 20% in week 24. The contribution of inorganic seston to total P was initially nearly 20% and decreased to lower proportions over the summer.

In the eutrophic lake (Fig. 5, right panels), phytoplankton N and P accounted for much greater proportions of TN and TP. The phytoplankton proportion of TN and TP increased and decreased with variations in Chl concentration, indicating that the concentration of other seston components were much less variable than the phytoplankton component. As observed with the oligo/mesotrophic lake, dissolved N still accounted for a large proportion of observed TN and dissolved P, a smaller proportion of observed TP.

DIN in 56 of 241 samples was $>50 \mu\text{g/L}$. Estimates of N and P concentrations per unit of Chl differed

Table 2. Estimates of coefficients for exploratory model for low and high DIN samples. Mean coefficient values shown with 90% credible intervals in parentheses.

Model	Coefficient	Low DIN	High DIN
TP	d_1	0.58 (0.51, 0.65)	0.73 (0.58, 0.89)
	d_2	7.1 (6.1, 8.0)	6.4 (4.4, 8.5)
TN	f_1	5.6 (4.8, 6.4)	4.0 (2.0, 6.0)
	f_2	71 (63, 81)	61 (39, 84)

substantially between low and high DIN samples, but because a smaller sample size yielded a greater uncertainty for estimates for the high DIN coefficients, none of the differences was statistically significant (Table 2). Mean P content associated with phytoplankton (d_1) increased from low to high DIN samples, whereas N content decreased (f_1). Both P and N content in VSS_{np} decreased slightly from low to high DIN samples (d_2 and f_2). These changes translated into a shift in N:P in phytoplankton from 21 (90% credible intervals: 18–25) to 12 (6–19) from low DIN to high DIN samples, while N:P in VSS_{np} exhibited a negligible change from 22 (19–27) to 21 (12–33).

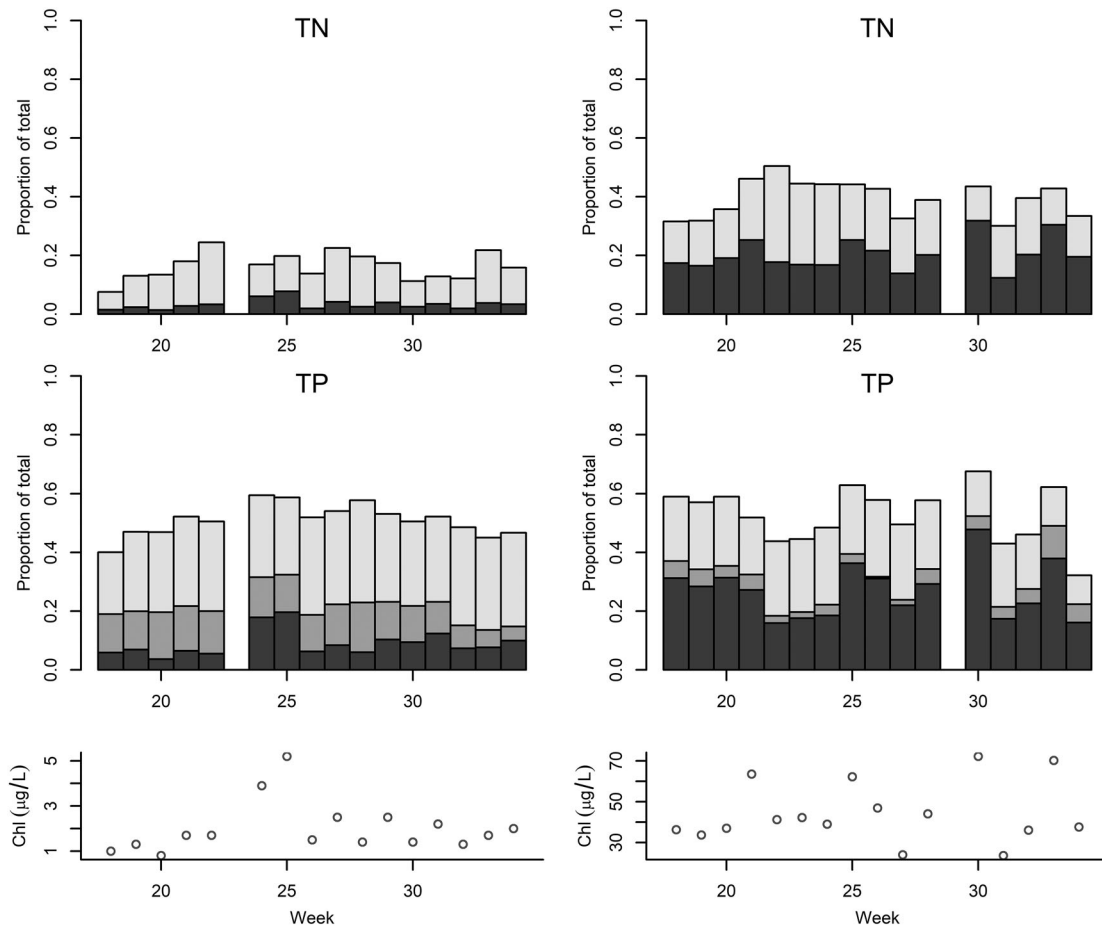


Figure 5. Allocation of P and N to different compartments for a mesotrophic lake (left column) and eutrophic lake (right column). Dark gray: proportion of nutrient attributed to phytoplankton biomass; medium gray: proportion of nutrient attributed to NVSS (TP plots only); and light gray: proportion of nutrient attributed to non-phytoplankton VSS. Dissolved proportions of N and P account for remaining proportions for each stack of bars.

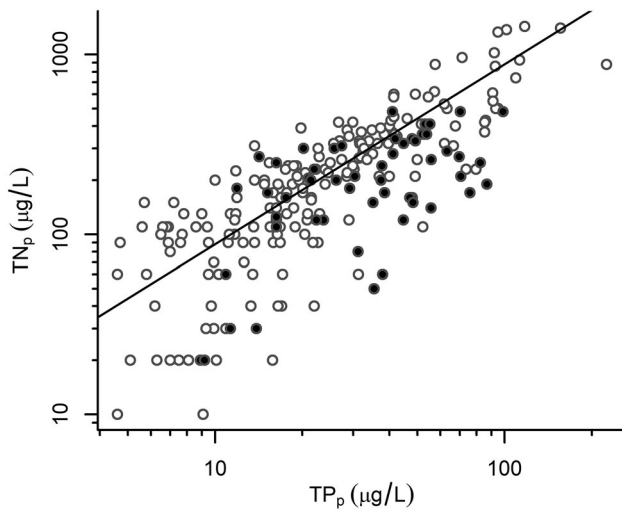


Figure 6. Relationship between TP_p and TN_p . Solid line: N:P = 19 ratio expected in phytoplankton dominated sample. Filled circles: samples with DIN > 50 $\mu\text{g/L}$.

Whole seston estimates of N:P (quantified as TN_p : TP_p) clustered around a value N:P = 19 estimated for phytoplankton (Fig. 6). N:P for VSS_{np} was slightly higher, and therefore an increase in the proportion of VSS_{np} in a sample would be expected to shift observations slightly above the solid line. Some of the large deviations of values below N:P = 19 can likely be attributed to samples in which nutrients are replete because phytoplankton in these samples may have much lower N:P (discussed earlier).

The limiting relationships between TP, TN, and Chl estimated using Missouri data were similar to the qualitative limits observed in the national dataset (Fig. 7). The deviation of the cloud of observations for TN from the limiting relationships can likely be attributed to the fact that only TN (rather than TN_p) was available from

the NLA dataset, and so observed TN is likely higher than TN_p for all samples because of the contribution from dissolved organic and inorganic N. The limiting relationship for TP more closely corresponded to the lower limits of the distribution of points, suggesting that the dissolved component of P accounted for a smaller proportion of TP. Overall, these results suggest that broadly applicable limiting relationships between TP, TN, and Chl can be estimated.

Discussion

We introduced a new approach for estimating stoichiometric ratios of different components of seston in lakes and reservoirs. The model accurately predicts variations in TP and TN among different samples and yields coefficients directly interpretable in terms of the nutrient content associated with different components of seston. Application of this modeling approach can help refine our understanding of the mechanisms that determine lake seston stoichiometry and inform the development of models used to predict Chl from nutrient inputs.

Bayesian network models provided a means of modeling relationships between measurements of different parameters in a single, internally consistent framework. The Bayesian model described here builds on statistical approaches described in earlier studies. In particular, our current formulation is similar to a statistical approach for separating seston into living and non-living components described by Hessen et al. (2003). The current approach expands that model to estimate separate contributions from $NVSS$ and VSS_{np} and uses gamma distributions to account for the increase in the variance of observations with mean values. Our approach is also similar to a technique for decomposing the contributions from organic and inorganic seston components

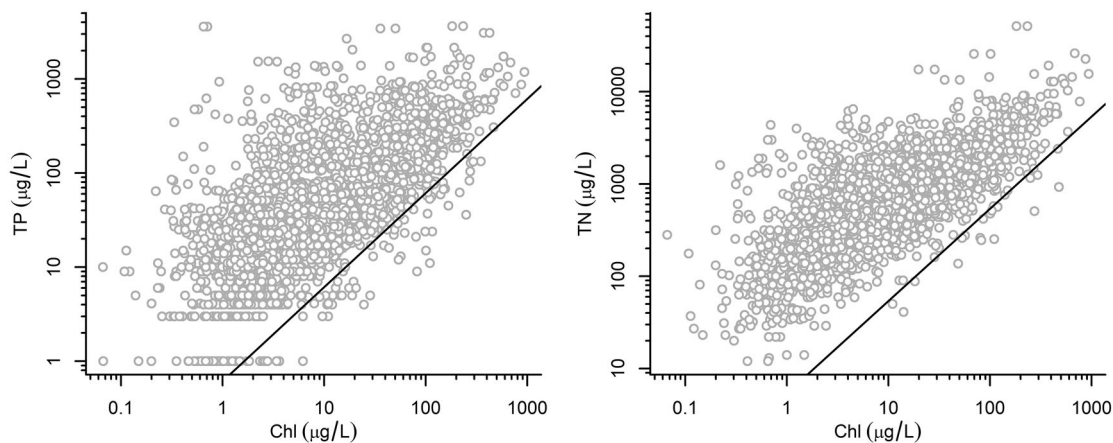


Figure 7. National lakes assessment observations of TP, TN, and Chl. Solid lines are relationships for TP_p and TN_p estimated from Missouri lakes.

described by Effler et al. (2014), who estimated the nutrient content of different components by analyzing a small number of samples dominated by only one constituent. Our estimates of limiting behavior are more precise than their approach because we use all the data and a model formulation that explicitly includes the limiting relationships that occur when one component of seston is the dominant contributor (Fig. 3).

Seston stoichiometry

Our analysis highlights how the amounts of N and P associated with different components of seston (phytoplankton, non-living organic material, and inorganic suspended sediment) contribute to variability in whole seston measurements of N:P. We found that N:P in organic non-living seston was only slightly higher than that estimated in phytoplankton. By contrast, N content in inorganic seston was negligible (Vitousek and Howarth 1991), and although P was present, as a proportion of total mass the P content of inorganic seston was much lower than in organic seston. In samples with large concentrations of inorganic sediment, however, inorganic seston can contribute to variations in whole seston P (Effler et al. 2014). Over the summer sampling season, the contributions of different seston components to TN and TP varied considerably (Fig. 5), manifested as variations in whole seston N:P.

Our findings that N:P = 19 and C:P = 170 in phytoplankton in our base model were strikingly similar to values estimated from analysis of whole seston from a global dataset of lakes (Hecky et al. 1993). In this earlier work, data were confined to lakes with residence times exceeding 6 months to limit the contributions of allochthonous seston, and the final values were estimated by regression of mean N and C content on mean P content across 44 lakes. Our present data were collected from reservoirs with retention times as low as 0.06 years, and we relied on modeling to control for the effects of inorganic and non-living organic matter. The similarity of the ratio estimates from these 2 contrasting datasets and analyses may suggest that mean nutrient content in the phytoplankton assemblages of different lakes converges to a single value, but analysis of data collected from more systems is required to solidify this conclusion.

Several mechanisms have been proposed that cause variations in whole seston N:P in addition to the differences among seston components on which our analysis is focused. Examples of these mechanisms include nutrient availability (Hecky et al. 1993), light to nutrient ratio (Sterner et al. 1997), temperature, water residence time (They et al. 2017), variations in nutrient content within different phytoplankton species, and variations in

phytoplankton species composition (Martiny et al. 2013). Our model does not exclude consideration of these other mechanisms. Instead, we suggest that examining the effects of these different mechanisms on particular components of seston via analyses similar to that described here can deepen our understanding of why N:P varies.

Our exploratory analysis of the effects of nutrient availability on N:P provides one example of how partitioning P and N among seston components may help clarify how environmental conditions affect seston stoichiometry. We found that N:P shifted dramatically from 21 for low DIN samples to 12 for high DIN samples. This change is qualitatively consistent with predictions of changes in N:P from a state of competitive equilibrium (i.e., low levels of biologically available nutrients) to a state of exponential growth (i.e., nutrient replete environments; Klausmeier et al. 2004a). Our exploratory analysis is only illustrative, however, and not intended as an exhaustive test of the effects of excess nutrients. Coupling our analysis of field data with manipulative experiments would provide more robust inferences. The type of analysis we describe here may then help bridge the gap between small-scale experiments and observations gleaned from large-scale monitoring data by providing a means of controlling for some of covariates inherent in field data (Elser and Hamilton 2007). The structure of our Bayesian model allows the introduction of variations in coefficient values, as in our exploratory analysis. In cases in which several groups of samples can be defined, coefficients can be expressed as a distribution of values characterized by a mean and variance (Gelman and Hill 2007) and a hierarchical structure imposed on coefficient values. We used this approach in our base model to estimate lake-specific mean values of inferred VSS_{np} , which improved our model for VSS. Ultimately, modeling the variations in nutrient content for different seston components would improve understanding of which variations in whole seston N:P can be attributed to differences in the composition of the seston versus differences in the nutrient content of different components.

Analyzing differences in the composition of seston components could also improve our understanding of the effects of variations in external loading of organic material and nutrients. For example, differences in watershed land use (Arbuckle and Downing 2001) and in the sources of nutrient loading (Downing and McCauley 2003) have been associated with changes in N:P observed in lakes. Similarly, atmospheric loading (Elser et al. 2009) has altered lake N:P, and particulate N:P varies with residence time (They et al. 2017). In these examples, examining the allocation of nutrients to different

seston components (Fig. 5) would likely advance understanding the effects of different types of nutrient loading on seston components.

Predicting Chl from nutrient concentrations

Statistical estimates of relationships between concentrations of TP and Chl and between TN and Chl in lakes have been fundamental to understanding and managing lake eutrophication (Jones and Bachmann 1976, Jones and Knowlton 2005), and our current models provide relationships between this same set of variables. Our model differs, however, from traditional nutrient/Chl models in 2 important ways that reflect our focus on the elemental composition of seston (Lewis and Wurtsbaugh 2008). First, we used covariates (NVSS and VSS_{np}) that accounted for variations in TP and TN, rather than in Chl. By contrast, most other analyses of nutrient/Chl relationships have sought to explain deviations in Chl concentrations from predictions using TP and/or TN. This interpretation of the nutrient/Chl relationship has led to investigations of the effects of covariates such as TN:TP (Guildford and Hecky 2000), zooplankton abundance (Mazumder 1994), lake morphology (Hoyer and Jones 1983, Phillips et al. 2008), and water transparency (Jones and Knowlton 2005, Webster et al. 2008).

We believe our model more accurately represents the processes that govern relationship between simultaneous observations of TP, TN, and Chl and provides the means to identify more specific hypotheses regarding the effects of other factors on nutrient/Chl relationships. For example, instead of simply hypothesizing that the presence of a high abundance of zooplankton grazers reduces Chl relative to TP, we can identify particular conditions in which zooplankton would be expected to have the strongest effects on the ratio of Chl to nutrients. In general, as zooplankton reduce phytoplankton abundance, the decrease in phytoplankton (as measured by Chl) should be accompanied by a decrease in the N and P associated with the phytoplankton component of seston. Other components of seston that contribute to TN and TP should be unchanged, so the effects of zooplankton should track lines associated with nearly constant values of the other seston components (as plotted in Fig. 3 and plotted with x and y variables reversed in Fig. 8). Then, in samples in which phytoplankton is the dominant source of particulate N and P, we would expect zooplankton grazing to reduce Chl, TN_p , and TP_p in the nearly the same proportions (e.g., change from A to B in Fig. 8), and the observed ratio between Chl and TN_p and between Chl and TP_p would be nearly constant. In samples in which phytoplankton is not the dominant source of particulate N and P, however, we would expect that

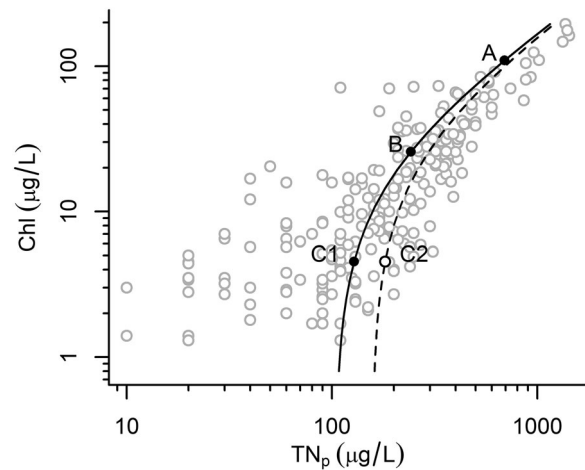


Figure 8. Hypothesized effect of zooplankton grazing on relationship between TN_p and Chl. Solid line: predicted relationship between TN_p and Chl assuming a constant value of VSS_{np} at the median value of VSS_{np} . Dashed line: predicted relationship between TN_p and Chl, assuming VSS_{np} equal to the 75th percentile of observed values.

with higher levels of grazing, the ratio between Chl and nutrients would decrease (a change from B to C1, Fig. 8). This effect may be magnified if zooplankton excretion increases the concentration of non-living organic seston (a change from B to C2, Fig. 8; Urabe et al. 2003). Hence, high abundance of zooplankton would be expected to reduce Chl concentration relative to nutrient concentrations, but this effect would be most pronounced in lakes in which phytoplankton is not the dominant component of seston. Hypotheses for the effect of other factors on TP–Chl and TN–Chl relationships could be developed in a similar manner.

The second way our current model differs from traditional investigations of nutrient/Chl relationships is that its mathematical formulation is based on untransformed measurements. By contrast, traditional nutrient/Chl models have been based on log-transformed variables. For example, if we apply the conventional approach to the model for TP, the following model equation results:

$$\begin{aligned} \log(\text{TP}) = & b_1 + P_{\text{diss}} + b_2 \log(\text{Chl}) \\ & + b_3 \log(\text{NVSS}) + b_4 \log(\text{VSS}_{np}) \\ & + \varepsilon_4, \end{aligned} \quad (10)$$

In this more commonly used model equation, the model coefficients become exponents of the predictor variables when the equation is expressed in original units and TP is modeled as the product rather than the sum of the different predictors. In this latter formulation, the model coefficients are not immediately interpretable in terms of nutrient content. Furthermore, coefficients are linear, and multiple linear regression can be used to

estimate their values, but the limiting behavior discussed earlier when a particular seston component dominates does not occur. Identifying these limits in the relationships between TP, TN, and Chl may be particularly important in broadening the applicability of models predicting Chl from nutrients. Indeed, our finding that the limiting relationships estimated in Missouri were qualitatively similar to the limits of observations collected in a continental scale dataset suggests that, at least at large spatial scales, assuming constant values for phytoplankton P and N content provides a useful approximation.

Uncertainties and conclusions

A few aspects of the current analysis may have introduced uncertainty and error into the results. We used Chl as a proxy for phytoplankton biomass, so any seston components that covaried with Chl would be treated in the analysis as if they were live phytoplankton. For example, nutrients associated with autotrophic organic detritus that increases with phytoplankton biomass would be indistinguishable from phytoplankton biomass. Further comparisons between the present analytical approach and direct measurements of different seston components would help quantify the magnitude of this error. Chl content in phytoplankton also can vary among species and with changes in environmental conditions (Hecky et al. 1993, Kasprzak et al. 2008), and estimating nutrient content in different types of lakes and different conditions will help quantify the magnitude of this variability.

Estimates of C content introduce another source of uncertainty to the current estimates of seston stoichiometry. We estimated C content as a fixed fraction of volatile suspended sediment, but this fraction can vary with conditions (Grove and Bilotta 2014). Others have employed different approaches for measuring organic C, including using a relationship between phytoplankton cell volume and C content (Hessen et al. 2003) and direct measurements of elemental C (Urbansky 2001). Concurrent measurements of VSS and direct measurements of elemental C would provide the most robust test of the method used here to estimate C.

Measuring dissolved and particulate fractions of N and P requires an arbitrary decision regarding the pore size of the filters because the dissolved fraction is defined as any material that passes through the filter (Solórzano and Sharp 2003). Similarly, suspended sediment and Chl are measured from material retained by the filters. Hence, the current findings with regard to the nutrient content of different seston components may change with data collected with finer or coarser filters, or with more involved measurements of dissolved

nutrients (Karl and Björkman 2015). Indeed, nephelometric measurements of the turbidity of filtrate in the present data suggest that a substantial amount of suspended sediment passes through the 1.5 μm filter (Knowlton and Jones 2000). Because data used here were internally consistent, such that dissolved and particulate nutrients, suspended sediment, and Chl were all measured using filters with comparable pore sizes, the estimates of nutrient content of various seston components are robust. Comparisons of the present results with other studies, however, should take into account the possible effect of different filter pore sizes.

The current statistical approach can be applied to measurements easily acquired in routine monitoring, a feature that allows broad use in analyses of both field observations and data from experimental manipulations. Collecting auxiliary information regarding the stoichiometry of different seston components could also enhance interpretation of the model results. For example, laboratory methods for separating different components (Hamilton et al. 2005, Efler et al. 2014) yield complementary data that, if conducted simultaneously with the current analysis approach, would provide data to further validate the approach.

In conclusion, we showed that we can accurately model TP and TN concentrations in lakes as a function of Chl, suspended sediment, and dissolved nutrients. These models provide insight into the nutrient content of different components of seston, which in turn can be interpreted with regard to the values of different environmental factors. The present models also inform efforts to develop eutrophication models that predict Chl as a function of nutrient loads.

Acknowledgements

We thank Daniel Obrecht, Jennifer Graham, Carol Pollard, and others for field collection and laboratory analyses. We also thank 2 anonymous reviewers whose comments improved this manuscript. The views expressed in this paper are those of the authors and do not reflect the official policy of the US Environmental Protection Agency. Support for this project came from the Missouri Department of Natural Resources, Missouri Agricultural Experiment Station, and the Food & Agriculture Policy Research Institute.

ORCID

Lester L. Yuan  <http://orcid.org/0000-0002-9462-824X>

References

Arbuckle KE, Downing JA. 2001. The influence of watershed land use on lake N:P in a predominantly agricultural landscape. *Limnol Oceanogr.* 46(4):970–975.

- Cotner JB, Hall EK, Scott T, Haldal M. 2010. Freshwater bacteria are stoichiometrically flexible with a nutrient composition similar to seston. *Front Microbiol.* 1:132.
- Dodds WK. 2006. Nutrients and the “dead zone”: the link between nutrient ratios and dissolved oxygen in the northern Gulf of Mexico. *Front Ecol Environ.* 4(4):211–217.
- Downing JA, McCauley E. 2003. The nitrogen:phosphorus relationship in lakes. *Limnol Oceanogr.* 37(5):936–945.
- Effler SW, Prestigiacomo AR, Peng F, Gelda R, Matthews DA. 2014. Partitioning the contributions of minerogenic particles and bioeston to particulate phosphorus and turbidity. *Inland Waters.* 4(2):179–192.
- Elser JJ, Andersen T, Baron JS, Bergstrom A-K, Jansson M, Kyle M, Nydick KR, Steger L, Hessen DO. 2009. Shifts in lake N:P stoichiometry and nutrient limitation driven by atmospheric nitrogen deposition. *Science.* 326(5954):835–837.
- Elser JJ, Hamilton A. 2007. Stoichiometry and the new biology: the future is now. *PLOS Biol.* 5(7):e181.
- Elser JJ, Hassett RP. 1994. A stoichiometric analysis of the zooplankton–phytoplankton interaction in marine and freshwater ecosystems. *Nature.* 370(6486):211–213.
- Elser JJ, Urabe J. 1999. The stoichiometry of consumer-driven nutrient recycling: theory, observations, and consequences. *Ecology.* 80(3):735–751.
- Findlay S, Pace M, Lints D. 1991. Variability and transport of suspended sediment, particulate and dissolved organic carbon in the tidal freshwater Hudson River. *Biogeochemistry.* 12(3):149–169.
- Gelman A, Hill J. 2007. Data analysis using regression and multilevel/hierarchical models. New York (NY): Cambridge University Press.
- Glibert PM, Fullerton D, Burkholder JM, Cornwell JC, Kana TM. 2011. Ecological stoichiometry, biogeochemical cycling, invasive species, and aquatic food webs: San Francisco estuary and comparative systems. *Rev Fish Sci.* 19(4):358–417.
- Grove MK, Bilotta GS. 2014. On the use of loss-on-ignition techniques to quantify fluvial particulate organic carbon. *Earth Surf Process Landf.* 39(9):1146–1152.
- Guildford SJ, Hecky RE. 2000. Total nitrogen, total phosphorus, and nutrient limitation in lakes and oceans: is there a common relationship? *Limnol Oceanogr.* 45(6):1213–1223.
- Hamilton SK, Sippel SJ, Bunn SE. 2005. Separation of algae from detritus for stable isotope or ecological stoichiometry studies using density fractionation in colloidal silica. *Limnol Oceanogr-Meth.* 3(3):149–157.
- Hassett RP, Cardinale B, Stabler LB, Elser JJ. 1997. Ecological stoichiometry of N and P in pelagic ecosystems: comparison of lakes and oceans with emphasis on the zooplankton–phytoplankton interaction. *Limnol Oceanogr.* 42(4):648–662.
- Hecky RE, Campbell P, Hendzel LL. 1993. The stoichiometry of carbon, nitrogen, and phosphorus in particulate matter of lakes and oceans. *Limnol Oceanogr.* 38(4):709–724.
- Hessen DO, Andersen T, Brettum P, Faafeng BA. 2003. Phytoplankton contribution to sestonic mass and elemental ratios in lakes: implications for zooplankton nutrition. *Limnol Oceanogr.* 48(3):1289–1296.
- Hillebrand H, Steinert G, Boersma M, Malzahn A, Meunier CL, Plum C, Ptacnik R. 2013. Goldman revisited: faster-growing phytoplankton has lower N:P and lower stoichiometric flexibility. *Limnol Oceanogr.* 58(6):2076–2088.
- Hoyer MV, Jones JR. 1983. Factors affecting the relation between phosphorus and chlorophyll *a* in Midwestern reservoirs. *Can J Fish Aquat Sci.* 40(2):192–199.
- Jones JR, Bachmann RW. 1976. Prediction of phosphorus and chlorophyll Levels in lakes. *J Water Pollut Control Fed.* 48(9):2176–2182.
- Jones JR, Knowlton MF. 2005. Chlorophyll response to nutrients and non-algal seston in Missouri reservoirs and oxbow lakes. *Lake Reserv Manage.* 21(3):361–371.
- Jones JR, Obrecht DV, Perkins BD, Knowlton MF, Thorpe AP, Watanabe S, Bacon RR. 2008. Nutrients, seston, and transparency of Missouri reservoirs and oxbow lakes: an analysis of regional limnology. *Lake Reserv Manage.* 24(2):155–180.
- Karl DM, Björkman KM. 2015. Chapter 5 - Dynamics of dissolved organic phosphorus. In: Hansell DA, Carlson CA, editors. *Biogeochemistry of marine dissolved organic matter* (2nd ed.). Boston (MA): Academic Press; p. 233–334.
- Kasprzak P, Padiśák J, Koschel R, Krienitz L, Gervais F. 2008. Chlorophyll *a* concentration across a trophic gradient of lakes: an estimator of phytoplankton biomass? *Limnologica.* 38(3–4):327–338.
- Klausmeier CA, Litchman E, Daufresne T, Levin SA. 2004a. Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton. *Nature.* 429(6988):171.
- Klausmeier CA, Litchman E, Levin SA. 2004b. Phytoplankton growth and stoichiometry under multiple nutrient limitation. *Limnol Oceanogr.* 49(4):1463–1470.
- Knowlton MF, Jones JR. 2000. Non-algal seston, light, nutrients and chlorophyll in Missouri reservoirs. *Lake Reserv Manag.* 16(4):322–332.
- Lewis WM, Wurtsbaugh WA. 2008. Control of lacustrine phytoplankton by nutrients: erosion of the phosphorus paradigm. *Int Rev Hydrobiol.* 93(4–5):446–465.
- Martiny AC, Pham CTA, Primeau FW, Vrugt JA, Moore JK, Levin SA, Lomas MW. 2013. Strong latitudinal patterns in the elemental ratios of marine plankton and organic matter. *Nat Geosci.* 6:279–283.
- Mazumder A. 1994. Phosphorus–chlorophyll relationships under contrasting herbivory and thermal stratification: predictions and patterns. *Can J Fish Aquat Sci.* 51(2):390–400.
- Perhar G, Arhonditsis GB. 2009. The effects of seston food quality on planktonic food web patterns. *Ecol Model.* 220(6):805–820.
- Phillips G, Pietiläinen O-P, Carvalho L, Solimini A, Lyche Solheim A, Cardoso A. 2008. Chlorophyll–nutrient relationships of different lake types using a large European dataset. *Aquat Ecol.* 42(2):213–226.
- Pribyl DW. 2010. A critical review of the conventional SOC to SOM conversion factor. *Geoderma.* 156(3):75–83.
- Ptacnik R, Andersen T, Tamminen T. 2010. Performance of the Redfield ratio and a family of nutrient limitation indicators as thresholds for phytoplankton N vs. P limitation. *Ecosystems.* 13(8):1201–1214.
- R Core Team. 2017. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.
- Redfield AC. 1958. The biological control of chemical factors in the environment. *Am Sci.* 46(3):230A–221.
- Solórzano L, Sharp JH. 2003. Determination of total dissolved phosphorus and particulate phosphorus in natural waters. *Limnol Oceanogr.* 25(4):754–758.

- Stan Development Team. 2016. Stan modeling language users guide and reference manual, version 2.14.0. <https://mc-stan.org>
- Sterner RW, Andersen T, Elser JJ, Hessen DO, Hood JM, McCauley E, Urabe J. 2008. Scale-dependent carbon:nitrogen:phosphorus seston stoichiometry in marine and freshwaters. *Limnol Oceanogr.* 53(3):1169–1180.
- Sterner RW, Elser JJ, Fee EJ, Guildford SJ, Chrzanowski TH. 1997. The light:nutrient ratio in lakes: the balance of energy and materials affects ecosystem structure and process. *Am Nat.* 150(6):663–684.
- Sterner RW, Hessen DO. 1994. Algal nutrient limitation and the nutrition of aquatic herbivores. *Annu Rev Ecol Syst.* 25(1):1–29.
- They NH, Amado AM, Cotner JB. 2017. Redfield ratios in inland waters: higher biological control of C:N:P ratios in tropical semi-arid high water residence time lakes. *Front Microbiol.* 8:1505.
- Urabe J, Nakanishi M, Kawabata K. 2003. Contribution of metazoan plankton to the cycling of nitrogen and phosphorus in Lake Biwa. *Limnol Oceanogr.* 40(2):232–241.
- Urbansky ET. 2001. Total organic carbon analyzers as tools for measuring carbonaceous matter in natural waters. *J Environ Monit.* 3(1):102–112.
- [US EPA] US Environmental Protection Agency. 2010. National lakes assessment: a collaborative survey of the nation's lakes. Washington (DC): Office of Water and Office of Research and Development Report No.: EPA 841-R-09-001.
- [US EPA] US Environmental Protection Agency. 2012. 2012 National lakes assessment. laboratory operations manual. Washington (DC): Report No.: EPA-841-B-11-004.
- [US EPA] US Environmental Protection Agency. 2017. National lakes assessment 2012: technical report. Washington (DC): US EPA, Office of Water, Office of Research and Development Report No.: EPA 841-R-16-114.
- Vitousek P, Howarth R. 1991. Nitrogen limitation on land and in the sea: how can it occur? *Biogeochemistry.* 13:87–115.
- Walling DE, Moorehead PW. 1989. The particle size characteristics of fluvial suspended sediment: an overview. *Hydrobiology.* 176-177:125–149.
- Webster KE, Soranno PA, Cheruvilil KS, Bremigan MT, Downing JA, Vaux PD, Asplund TR, Bacon LC, Connor J. 2008. An empirical evaluation of the nutrient-color paradigm for lakes. *Limnol Oceanogr.* 53(3):1137–1148.