Rethinking phosphorus–chlorophyll relationships in lakes

Lester L. Yuan [1]*, **John R. Jones** [2]
[1]Office of Water, U.S. Environmental Protection Agency, Washington, District of Columbia
[2]School of Natural Resources, University of Missouri, Columbia, Missouri

Abstract

The empirical relationship between total phosphorus and chlorophyll has guided lake management decisions for decades, but imprecision in this relationship in individual lakes limits the utility of these models. Many environmental factors that potentially affect the total phosphorus-chlorophyll relationship have been studied, but here we hypothesize that imprecision can be reduced by considering differences in the proportions of phosphorus bound to three different "compartments" in the water column: phosphorus bound in phytoplankton, phosphorus bound to suspended sediment that is not associated with phytoplankton, and dissolved phosphorus. We specify a hierarchical Bayesian network model that estimates phosphorus associated with each compartment using field measurements of chlorophyll, total suspended solids, and total phosphorus collected from reservoirs in Missouri, United States. We then demonstrate that accounting for these different compartments yields accurate predictions of total phosphorus in individual lakes. Results from this model also yield insights into the mechanisms by which lake morphometric and watershed characteristics affect observed relationships between total phosphorus and chlorophyll.

Empirically estimated relationships between total phosphorus (TP) and chlorophyll (Chl) have provided a basis for lake management for over four decades. This relationship was initially identified in Connecticut and Japanese lakes (Deevey 1940; Sakamoto 1966), and subsequently extended to a broad range of temperate lakes in the mid-1970s (Dillon and Rigler 1974; Jones and Bachmann 1976; Carlson 1977). These early analyses regressed Chl on TP and reported similar coefficients showing the ratio of Chl:TP increased with lake trophic state. Over time, scores of published relationships have explored the veracity of this relationship (McCauley et al. 1989; Prairie et al. 1989; Jones and Knowlton 2005; Filstrup et al. 2014), assessed sources of residual variation, and tested the limits of applicability to different regions and lake types. Variations in the relationship have been attributed to differences in lake depth (Pridmore et al. 1985), TN:TP ratio (Smith 1982; Prairie et al. 1989; Molot and Dillon 1991), grazing by zooplankton and mussels (Mazumder 1994; Mellina et al. 1995), landscape characteristics (Wagner et al. 2011), and light limitation (Hoyer and Jones 1983; Knowlton and Jones 2000; Havens and Nürnberg 2004). Regional studies have evaluated the relationship as influenced by edaphic and climatic factors in locations such as Canada (Prepas and Trew 1983), Argentina, (Quirós 1990), the United Kingdom (Spears et al. 2013), and Europe (Phillips et al. 2008). Recently, lake classifications have

improved the precision and accuracy of this relationship (Yuan and Pollard 2014).

When seasonal averages of TP and Chl are used in cross-system analysis over a broad trophic range, a relatively precise relationship is frequently observed (Knowlton et al. 1984; Jones et al. 1998). In these types of analysis, mean values of TP and Chl are computed, so the effects of within-lake variability are reduced, improving the precision of the estimated relationship. Relationships between TP and Chl within individual lakes, however, often exhibit substantial variability about the overall mean. Some of this variability can be attributed to using measurements from single samples to estimate these relationships (increasing the effects of sampling variability), while additional variability has been attributed to differences in lake characteristics that affect the efficiency with which phosphorus is converted to algal biomass (Smith and Shapiro 1981; Rast et al. 1983). In most management applications, accurate predictions of response in individual lakes are needed for supporting decisions, and so, variability in TP-Chl relationships among lakes limits the utility of these models.

A TP measurement is comprised of P contained within different compartments, including P bound in phytoplankton, P bound to suspended sediment, and dissolved P (i.e., chemically dissolved P and P bound to particles small enough to pass through a filter) (Effler and O'Donnell 2010). In many lakes much of measured TP is associated with phytoplankton, and so, differences in phytoplankton biomass among lakes would be associated with differences in both Chl

*Correspondence: yuan.lester@epa.gov

and TP, yielding a strong correlation between the two (Lewis and Wurtsbaugh 2008). In other lakes, high concentrations of suspended sediment contribute to TP and affect the observed TP–Chl relationships (Jones and Knowlton 2005). When estimating TP–Chl relationships, lakes with high concentrations of suspended sediment show low Chl:TP ratios relative to the average pattern (Hoyer and Jones 1983; Jones and Knowlton 2005).

Here, we describe an alternate approach for modeling the relationship between TP and Chl in which we explicitly model the contributions of different compartments to observed TP. In doing so, we reverse the positions of TP and Chl in the model equation, seeking to explain variations in TP in various compartments, rather than seeking to explain variation in Chl. We illustrate our modeling approach with data collected from Missouri (MO) reservoirs, and we hypothesize that this new model can better account for variability in TP–Chl relationships among lakes.

Data

Monitoring data collected by the University of Missouri in 155 reservoirs during summers 1989–2016 were used for this analysis (Fig. 1). Missouri reservoirs vary broadly in their characteristics, ranging from relatively clear conditions in the forested south to turbid conditions in the agricultural north (Jones et al. 2008*a*). Composited surface samples were collected from most reservoirs three to four times during May–August near the dam. Samples were transported on ice to a field laboratory and processed by a standard methodology (Knowlton and Jones 1995). Total suspended solids (TSS) were determined by filtering a known volume of lake water through a Whatman934-AH filter (nominal filter size: 1.5 μm) that was prerinsed, dried, ashed, and tared. Chl (uncorrected for degradation products) was measured from material retained on a
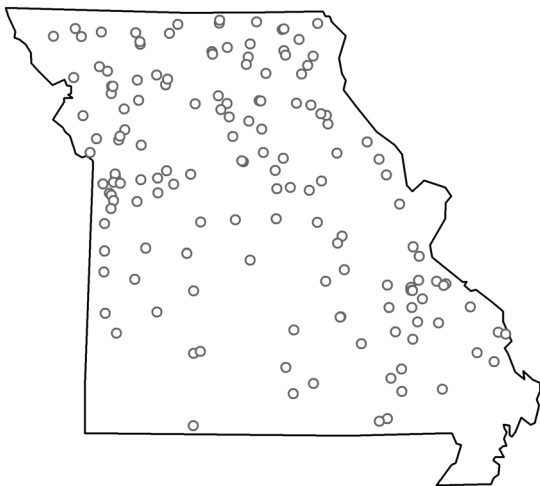
Fig. 1. Locations of sampled reservoirs.

1 μm Gelman AE filter, while TP was measured from the whole water sample.

In 2004, more intensive measurements were collected from 15 lakes. In these weekly samples, measurements of TP in filtrate (i.e., dissolved TP or dTP) were recorded. We used these measurements to test whether the model specified below accurately estimated dTP.

We also assembled watershed land use, quantified by the proportion of the upstream catchment used for row crop agriculture, and morphometric characteristics for each of the lakes in the data set. Lake morphometric data included lake mean depth, volume, flushing rate, and the ratio between lake surface area and catchment area.

At least 20 samples were available from each of the 155 reservoirs included in the data set, although in some locations, over 100 samples were available. The total number of samples was 7948.

Statistical analysis

We specified a model to estimate contributions to TP from three components: dissolved P, P bound to nonphytoplankton sediment, and P bound in phytoplankton. Direct measurements of nonphytoplankton sediment were not available. Instead, TSS was measured, which, like TP, includes contributions from both nonphytoplankton and phytoplankton components. So, modeling the network of the relationships among TSS, TP, and Chl is necessary to accurately estimate contributions from the three components of TP. A hierarchical structure is also specified in the model, such that different coefficient values are estimated for each reservoir in the data set, and hyper-distributions are used to specify the relationships among these coefficients. These hierarchical structures are described in detail below.

In the first relationship in the network, TSS was modeled as the sum of two components: (1) suspended sediment that is directly associated with phytoplankton biomass, or autochthonous suspended sediment (SS_{aut}) and (2) suspended sediment associated with all other sources (SS_{np}, or nonphytoplankton suspended sediment) (Fig. 2). This second component of suspended sediment includes sediment supplied by allochthonous sources and sediment resuspended from the lake basin (Hamilton and Mitchell 1996). We assumed that SS_{aut} is directly proportional to Chl (Jones et al. 2008*b*), a measure of algal biomass, and therefore, we expressed a model relationship for the components of TSS as follows:

$$TSS = SS_{np} + SS_{aut} = SS_{np} + bChl^k \qquad (1)$$

where we assumed that the amount of SS_{aut} associated with each unit of Chl varied with algal composition (Nalewajko 1966; Stabel 1986), which in turn, varied with trophic conditions (Godfrey 1982). Therefore, we expressed the second term of Eq. (1) as a power function of Chl.
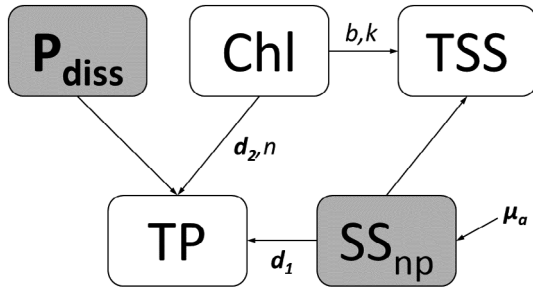
Fig. 2. Schematic representation of Bayesian network. P_{diss}: dissolved phosphorus, Chl: Chlorophyll concentration, TSS: total suspended solids, TP: total phosphorus, and SS_{np}: nonphytoplankton suspended solids. Shaded boxes indicate variables that are estimated in the model. Parameter labels next to arrows indicate coefficients estimated by the model (see text). Bolded variable (P_{diss}, μ_a, d_1, and d_2) indicate parameters for which lake-specific values are computed.

Measurements of TSS and Chl were highly skewed, and log transformations were required to effectively fit the model relationship to observed data:

$$\log(TSS_j) = \log\left(SS_{np,j} + bChl_j^k\right) + e_{TSS,j} \qquad (2)$$

where the subscript, j, refers to measurements in different samples, and $e_{TSS,j}$ is a normally distributed random error with a mean of zero and a standard deviation of σ_{TSS}. The coefficient, b, was assumed to be log-normally distributed to constrain it to positive values, while the exponent, k, was assumed to be normally distributed.

We next assumed that the mean concentrations of SS_{np} varied among lakes and that the concentration of SS_{np} estimated for each sample could be modeled as a log-normal distribution about the lake specific mean value of SS_{np}:

$$\log(SS_{np,j}) \sim N\left(\mu_{a,i[j]}, \sigma_a\right) \qquad (3)$$

where $\mu_{a,i}$ is the mean value of $\log(SS_{np})$ for lake i, corresponding to sample j, and σ_a is the standard deviation of the distribution of individual measurements of SS_{np}. The set of values for $\mu_{a,i}$ are then assumed to be drawn from a single normal distribution:

$$\mu_{a,i} \sim N\left(\mu, \sigma_\mu\right) \qquad (4)$$

where μ and σ_μ are the mean and standard deviation of this distribution. This overarching mean distribution loosely constrains the possible values of $\mu_{a,i}$, while allowing lakes with smaller amounts of data to "borrow information" from lakes with larger amounts of data (Gelman and Hill 2007).

Results from the model for TSS are used simultaneously to estimate contributions to different components of TP. Recall, we are modeling TP as being composed of contributions from dissolved P (P_{diss}), P that is bound to SS_{np}, and P that is bound

in phytoplankton. Based on this initial assumption, we can write the following model relationship:

$$TP = P_{diss} + d_1 SS_{np} + d_2 Chl^n \qquad (5)$$

where the concentration of P bound to non-phytoplankton suspended sediment is modeled as being directly proportional to SS_{np}. Similar to the model for TSS, we assumed that the quantity of P bound in phytoplankton changes with eutrophication status, and therefore model it as being proportional to a power function of Chl.

Log-transformations are required again to fit to observed data, and so we write the following expression:

$$\log(TP_j) = \log\left(P_{diss,i[j]} + d_{1,i[j]}SS_{np,j} + d_{2,i[j]}Chl_j^n\right) + e_{TP,j} \qquad (6)$$

where j indexes individual samples and $i[j]$ indexes different lakes associated with each sample. The random error $e_{TP,j}$ was assumed to be normally distributed with a mean of zero and a standard deviation of σ_{TP}.

We hypothesized that the coefficients d_1, d_2, and the magnitude of P_{diss} varied among lakes due to differences in catchment and lake characteristics, so we allowed different lake-specific values for each of these parameters. Overall, values for each parameter were assumed to be drawn from log-normal distributions to constrain them to positive values:

$$\log(P_{diss,i}) \sim N(\mu_{diss}, \sigma_{diss})$$
$$\log(d_{1,i}) \sim N(\mu_{d1}, \sigma_{d1}) \qquad (7)$$
$$\log(d_{2,i}) \sim N(\mu_{d2}, \sigma_{d2})$$

All of the relationships described earlier were fit simultaneously to the available data with a hierarchical Bayesian model (Stan Development Team 2016). Prior distributions for the hyper-parameters, μ, μ_{diss}, μ_{d1}, and μ_{d2} were specified as normal distributions with mean values of zero and standard deviations much greater than the expected value of the parameter to ensure that the prior distributions did not influence the results. Prior distributions for standard deviations, σ_μ, σ_{diss}, σ_{d1}, σ_{d2}, σ_{TSS}, and σ_{TP} were similarly noninformative, specified as half-Cauchy distributions with scale parameters much greater than the expected values of the parameters. Prior distributions for the two exponents, n and k, and for the coefficient, $log(b)$, were also noninformative normal distributions.

For comparison, we fit a simple linear regression model of the following form:

$$\log(TP_j) = c_1 + c_2 \log(Chl_j) + e_j \qquad (8)$$

where c_1 and c_2 are regression coefficients and e_j is a normally distributed residual error. The form of this equation is similar

to that commonly used to estimate relationships between TP and Chl except, to facilitate a comparison with the present model, the dependent variable in this formulation is TP.

We assessed the performance of the model by quantifying the error in model predictions of TP with mean square error (MSE), defined as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(\log\left(TP_{pred,i}\right)^2 - \log\left(TP_{obs,i}\right)^2\right) \qquad (9)$$

where N is the number of samples, TP_{pred} is the predicted TP concentration, and TP_{obs} is the observed TP concentration. Using this metric to examine prediction error allowed us to quantify the relative contribution of different lake-specific coefficients to the predictive performance of the model. More specifically, we computed several different predictions of TP as follows: (1) the simple linear regression model prediction of TP, (2) the full Bayesian network model prediction of TP, using lake-specific coefficients for P_{diss}, d_1, and d_2, (3) the "regional" Bayesian network model prediction of TP using the overall mean values for these same coefficients: μ_{diss}, μ_{d1}, and μ_{d2}, and (4) separate coefficient model predictions of TP, each using one of the lake-specific coefficients (P_{diss}, d_1, or d_2) and overall mean values for the other two. Based on the MSEs for each of the model predictions, we could estimate the proportional improvement in model accuracy associated with incorporating lake-specificity for each of the different model coefficients.

To help visualize the effects of modeling contributions of P_{diss} and SS_{np} to TP, we computed an "adjusted" value of TP (TP_{adj}) by subtracting the contributions of P_{diss} and SS_{np} from observed values of TP:

$$TP_{adj,j} = TP_j - P_{diss,i[j]} - d_{1,i[j]}SS_{np,j} \qquad (10)$$

We then compared plots of TP vs. Chl and TP_{adj} vs. Chl for different lakes.

Finally, we explored whether lake catchment characteristics and morphology were associated with the values of each of the lake-specific coefficients (P_{diss}, d_1, and d_2). Our intent for this data exploration was to identify potential linkages for future research, and to that end, we only computed Pearson correlation coefficients, and reported on watershed and lake morphological characteristics that were strongly correlated with the lake-specific coefficient values.

Results

Observations of TSS, Chl, and TP in the data set spanned a broad range of conditions (Table 1). TSS was correlated with Chl, and a distinct lower boundary in the scatter of data was evident (Fig. 3). The model relationship defining this lower boundary can be computed by setting SS_{np} to zero in Eq. (2). Then, after simplifying, we can write log(*TSS*) = log(*b*) + *k*log

(*Chl*). So, when SS_{np} is negligibly small, the relationship between SS_{aut} and Chl is a straight line in the plot of log(Chl) vs. log(TSS) (solid line in Fig. 3). Deviations in sampled values above this line show the contribution of SS_{np} to the overall TSS measurement. Small deviations below this line can be attributed to sampling variability of TSS and was estimated as $\sigma_{TSS} = 0.26$. Mean values of log(*b*) and *k* estimated from the model were − 0.51 (−0.56, −0.46) and 0.67 (0.66, 0.68) (90% credible intervals shown in parentheses). Based on the functional form we assumed for the relationship between TSS and Chl, we can infer that the contribution of phytoplankton to TSS (i.e., SS_{aut}/Chl) is proportional to $Chl^{-0.33}$. That is, as Chl increases, the amount of suspended sediment associated with each unit of Chl decreases, a trend which is consistent with a shift from diatom-dominated assemblages to green algae and cyanobacteria dominated assemblages (Nalewajko 1966).

Limiting relationships that estimate the P-content of phytoplankton biomass and SS_{np} can also be calculated (Fig. 4). For phytoplankton biomass, this limiting relationship is calculated by setting P_{diss} and SS_{np} in Eq. 6 to zero, yielding the following, log-transformed relationship: log(*TP*) = log(d_2) + *n* log (*Chl*). Different values of d_2 were estimated for each lake, but the distribution of these values is characterized by an overall mean and a standard deviation (Table 2). The untransformed mean value of d_2 among all lakes was 2.4 (2.0, 2.8), and the mean value of the parameter *n* was 0.69 (0.65, 0.72). The straight line based on these two parameter values represents P associated with phytoplankton biomass, as quantified by Chl, and it closely tracks the lower limit of the observed data (solid line, left panel Fig. 4). Again, small deviations below this line can be attributed to sampling variability of TP and was estimated as $\sigma_{TP} = 0.19$.

For SS_{np}, setting P_{diss} and Chl to zero yields the following relationship: log(*TP*) = log(d_1) + log(SS_{np}). The coefficient d_1 also varied among lakes, with an overall mean value of 5.5 (4.9, 6.1). In this case, the limiting relationship corresponds to a line with an intercept of log(5.5) and a slope of 1 (right panel, Fig. 4).

In untransformed units, values of P_{diss} ranged from 0.6 to 95.9 μg L^{-1} among the 155 lakes in the data set, while the P content of SS_{np} (d_1) ranged from 0.10% to 6.2% (Table 2). Mean values of SS_{np} (μ_a) ranged from 0.16 to 30.9 mg L^{-1}.

The P-content of autochthonous material is a function of both the coefficient d_2 and Chl. More specifically, we can express the P-content of autochthonous material as a Chl yield as follows,

$$\frac{Chl}{P} = \frac{Chl^{1-n}}{d_2}$$

Hence, Chl/P increases with Chl concentration at a rate proportional to Chl^{1-n}, or $Chl^{0.31}$ (*n* was previously estimated as 0.69). Because d_2 varies among lakes, the relationship between Chl/P and Chl also varies among lakes, but the lake-specific relationships generally clustered tightly about the

Table 1. Summary statistics of observed variables.

|	5th percentile	25th percentile	Median	75th percentile	95th percentile
Chl (μg L^{-1})	1.7	5.8	14.1	29.6	75.0
TP (μg L^{-1})	8	19	36	63	140
TSS (mg L^{-1})	1.4	3.4	6.1	10.0	19.8
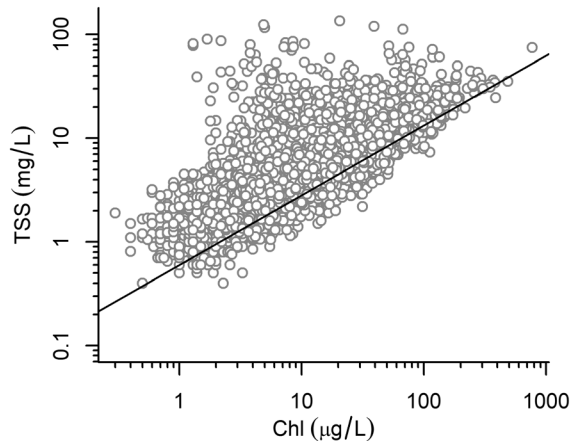
Fig. 3. TSS vs. Chl. Open circles: all sampled values, solid line: estimated suspended solids concentration associated with Chl calculated by setting contribution of SS_{np} to zero.

relationship based on overall mean values (Fig. 5). For comparison, Jones and Bachmann (1976) estimated log(Chl) = −1.09 + 1.46 log(TP), and to plot this relationship on the same axes, it can be reexpressed as Chl/TP = 0.18 Chl$^{0.315}$. This relationship (dashed line in Fig. 5) is located below all of the lake-specific relationships estimated in the current analysis, a difference that stems from the fact that the Jones and Bachmann (1976) relationship is based on TP, which includes

contributions from both P_{diss} and P bound to SS_{np}, whereas the current relationship is an estimate only of P bound in phytoplankton. Hence, TP values in Jones and Bachmann (1976) are higher than those in the current relationship, and Chl/P is lower. The current estimate of the exponent on Chl and that of Jones and Bachmann (1976) are strikingly similar, though.

One assumption in this model is that differences in TP concentration not correlated with SS_{np} or Chl provided an estimate of P_{diss}. In the subset of lakes for which measurements of the dissolved fraction of TP (dTP) were recorded ($n = 15$), we compared these measurements with estimates of mean P_{diss} in each lake (left panel, Fig. 6). The estimated values of P_{diss} defined a lower bound for observations of dTP (dashed line in plot), but in many lakes observed dTP concentrations were substantially greater than the model estimate. Closer examination of the measurements of dTP in the lakes in which P_{diss} underestimated dTP indicated that measurements of dTP were strongly associated with SS_{np}. One example of this association shown in the right panel (Fig. 6).

The MSE of the full model prediction of log(TP), using lake-specific values for all the parameters in the model, was 0.02, whereas the MSE of the "regional" model in which we computed predictions for log(TP) using overall mean values for the coefficients (μ_{diss}, μ_{d1}, and μ_{d2}) was 0.15. The MSE based on a simple linear regression fit to the data was 0.29, so prediction error of the full Bayesian network model was only 7% of the
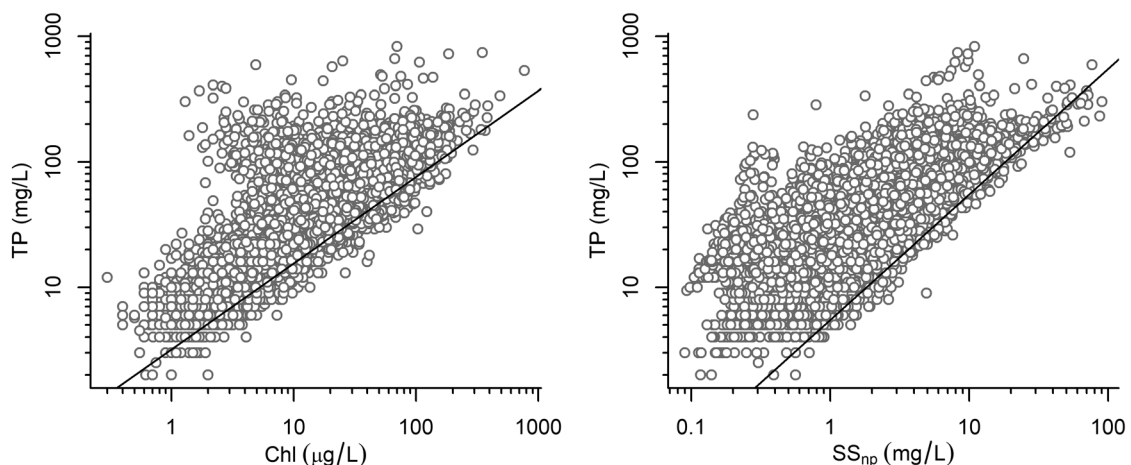
Fig. 4. Relationship between TP and different contributing components. Left panel: TP vs. Chl, right panel: TP vs. SS_{np}. Open circles: observed data. Solid line (left panel): estimated phosphorus concentration associated with Chl calculated by setting dissolved P and P associated with SS_{np} to zero. Solid line (right panel): estimated phosphorus concentration associated with SS_{np} calculated by setting P associated with Chl to zero.

Table 2. Mean values of parameters that defined the distributions of lake-specific model coefficients. Ninety percent credible intervals in parentheses.

	Mean	Standard deviation
μ_a	0.54 (0.40, 0.67)	0.95 (0.86, 1.06)
$\log(P_{diss})$	1.74 (1.46, 1.96)	1.05 (0.92, 1.22)
$\log(d_1)$	1.70 (1.59, 1.81)	0.67 (0.60, 0.76)
$\log(d_2)$	0.87 (0.71, 1.04)	0.27 (0.23, 0.31)
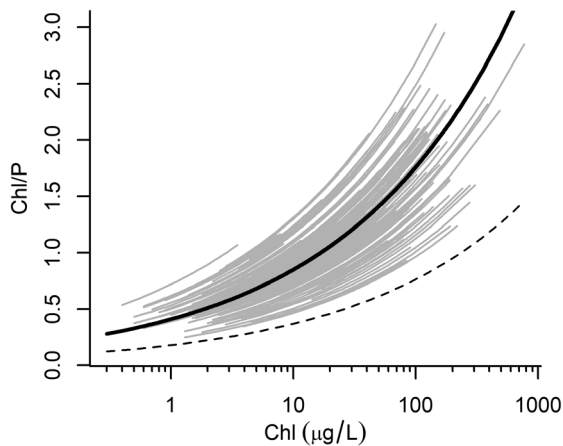
Fig. 5. Chlorophyll yield. Gray lines: relationships for individual lakes. Black line: overall mean relationship. Dashed line: Chl/TP ratio predicted by Jones and Bachmann (1976).

error of the simple linear regression. When lake-specific values for P_{diss} were used instead of μ_{diss} in the regional model, the predictions of $\log(TP)$ accounted for 34% of the difference in MSE between the full model and the regional model. Including lake-specific values for d_1 accounted for 45% of the

difference, while including lake-specific values for d_2 only accounted for 18% of the difference.

MSE estimates based on individual lakes mirrored trends observed using the full data set. When different simple linear regression models were fit to data from each lake, MSE values ranged from 0.02 to 0.52. MSE values for each lake could also be computed using the full Bayesian network model and lake-specific coefficient values and these ranged from 0.01 to 0.06. MSE values based on the Bayesian network model were less than those estimated by simple linear regression for all but three lakes, and on average, reduced the MSE by 70%.

In individual lakes, accounting for P_{diss} and SS_{np} markedly improved the qualitative strength of the association between TP and Chl. For a lake with relatively low concentrations of SS_{np} (left panel, Fig. 6), plotted values of unadjusted TP were only slightly greater than adjusted TP, but the variance in values about the mean relationship was reduced. In other lakes, SS_{np} concentrations were high and variable, and the effects on the TP–Chl relationship were strong. In the example lake shown (right panel, Fig. 7), raw measurements of TP and Chl exhibited a weak relationship, whereas adjusted TP was strongly associated with Chl.

Striking differences were observed in the amount of phosphorus associated with different compartments in these two example lakes. On average, in Sims Valley Lake, nearly 65% of phosphorus is bound in phytoplankton (Fig. 8), whereas in the entire data set, 46% of phosphorus is bound in phytoplankton. In contrast, in Manito Lake, over 75% of phosphorus in the samples collected were associated with SS_{np}, and only 18% was bound in phytoplankton.

Lake-specific values for P_{diss} were negatively associated with depth ($r = -0.48$) (Table 3), whereas the mean concentration of SS_{np} (μ_a) increased with percentage crops ($r = 0.44$). Correlations between d_1 (the P-content of SS_{np}) with different lake characteristics were all weak, with all correlation coefficients
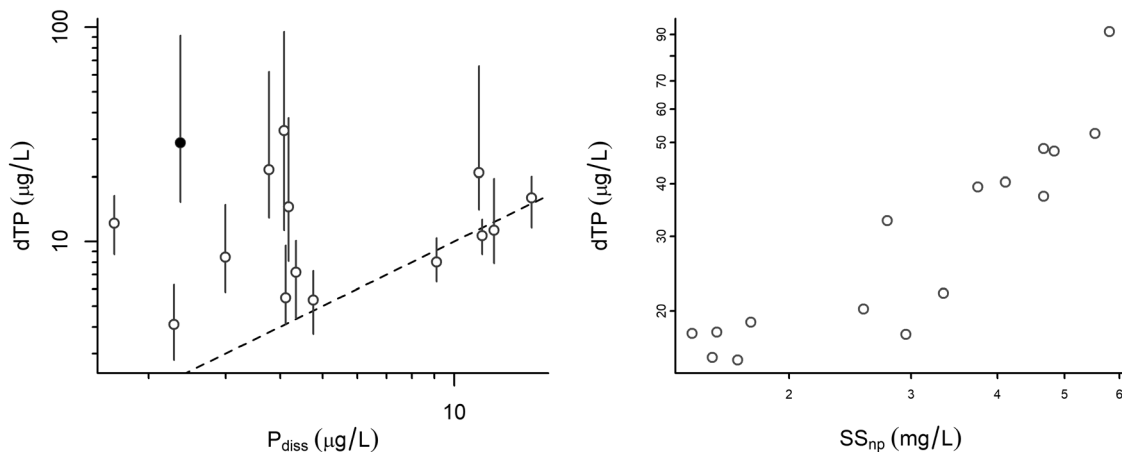
Fig. 6. Left panel: Comparison of estimated P_{diss} vs. observed dissolved total phosphorus (dTP). Vertical segments show the range of observed values of dTP for each lake. Open circle shows the mean dTP for each lake. Dashed line shows 1:1 relationship. Right panel: Relationship between individual measurements of dTP and SS_{np} for the site indicated in the left panel with a filled circle.
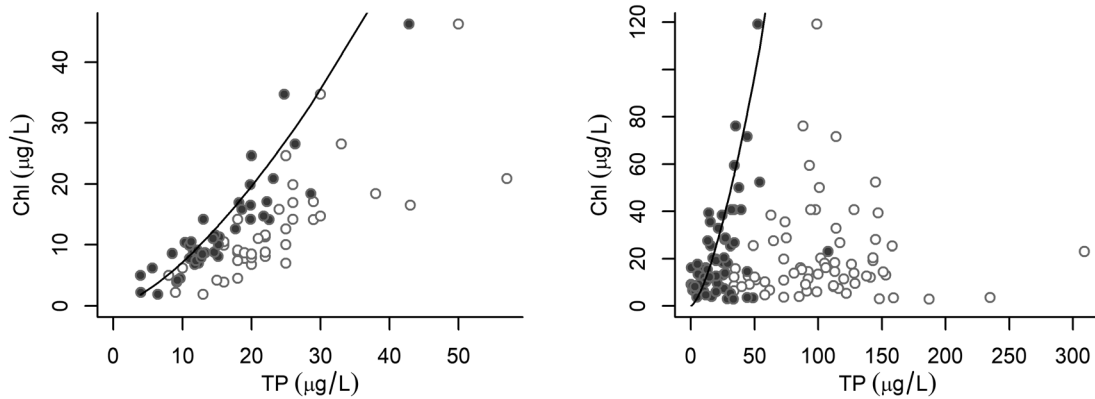
Fig. 7 Performance of TP-Chl model for two lakes (left panel: Sims Valley Lake, right panel: Manito Lake). Open circles: instantaneous samples of Chl and TP; filled circles: TP adjusted for P_{diss} and SS_{np} vs. observed Chl; solid line: predicted relationship between TP and Chl from Eq. (6).

less than 0.2. Lake-specific values of d_2 quantified differences among lakes in the P-content of autochthonous material beyond what could be attributed to differences in trophic status. Values of d_2 were weakly correlated with lake depth ($r = -0.37$).

Discussion

The new approach for modeling relationships between TP and Chl improves the accuracy of predictions and yields insights into the causes of variability in these relationships in individual lakes. The Bayesian network model simultaneously represents different relationships linking measurements of TSS, Chl, and TP, and by doing so, the model explicitly estimates

contributions from three different components of TP: P_{diss}, P bound to sediment, and P bound in phytoplankton.

In the first stage of the network model, we estimated components of TSS associated with phytoplankton vs. components of TSS associated with other sources (including allochthonous loads and resuspended sediment). A comparable laboratory approach to distinguish between sources of TSS is to measure nonvolatile and volatile components of sediment and to assume that the volatile fraction characterizes the contribution from phytoplankton (Knowlton and Jones 2000). However, samples in this study span the eutrophication gradient, and the shift along this gradient from diatom-dominated assemblages that are ~30–50% ash to green algae that are ~10% ash (Nalewajko 1966) would introduce errors in the laboratory approach. Assumptions underlying the statistical approach also may introduce errors. We assumed that, after accounting for eutrophication status and lake-specific characteristics, the proportion of TSS associated with each unit of Chl is fixed. In reality, this proportion can vary in time and space. However, the clearly defined lower boundary in the plotted relationship between Chl and TSS and the correspondence between the modeled limiting relationship and this boundary lends empirical support to the validity of our assumptions.

The current model increased the accuracy of predictions of TP based on Chl, addressing a long-standing issue with TP-Chl models in which relationships within individual lakes have
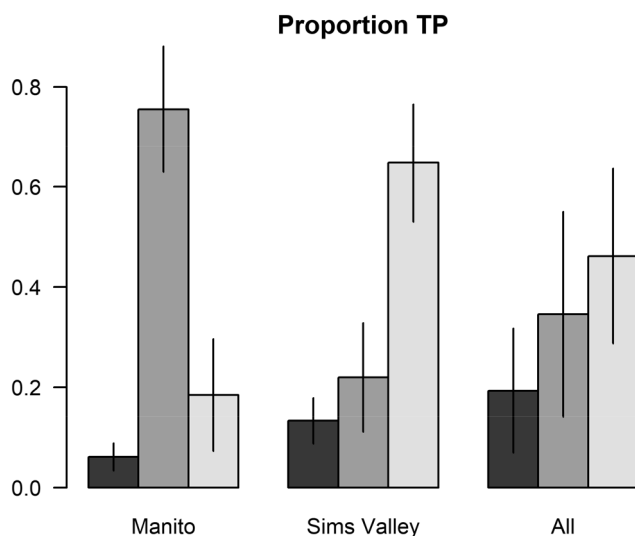
Fig. 8 Mean proportion of TP associated with different compartments for Manito, Sims Valley Lake, and all data. Leftmost dark gray segment: P_{diss}; middle gray segment: P bound to SS_{np}; rightmost light gray segment: P bound to phytoplankton. Vertical line segments show the standard deviation of proportions among samples collected from each lake.

Table 3. Correlation coefficients between lake-specific model parameters and lake morphological and watershed characteristics.

	Flush rate	Volume	Area ratio	Depth	Crops
μ_a	0.08	−0.09	0.08	−0.34	0.44
P_{diss}	0.11	−0.33	0.02	−0.48	0.39
d_1	−0.03	0.06	−0.05	0.03	0.18
d_2	0.26	−0.30	0.16	−0.37	0.14

varied substantially from a cross-system relationship based on mean values of TP and Chl from different lakes (Smith and Shapiro 1981; Spears et al. 2013). As described in the Introduction, previous studies have accounted for differences in observed Chl by examining covariates, whereas this analysis suggests that a primary source of variability in TP–Chl relationships derives from differences in SS_{np}, which alters the concentration of TP measured in a sample. Relationships based on measurements of TP and Chl in individual lakes benefit most from this adjustment because episodic loads of SS_{np} exert the strongest effects on individual measurements (Knowlton and Jones 1995). In contrast, seasonal averages of TP and Chl in a lake reduce the effect of SS_{np} to a single mean value. This mean contribution of SS_{np} then exerts a weaker influence on cross-system relationships estimated using seasonal averages, a phenomenon that may explain the similarity of the values of the exponent on Chl estimated from the current study and that estimated from analysis of seasonal mean TP and Chl (Jones and Bachmann 1976).

The present approach for modeling contributions to TP from different compartments refines our understanding of factors affecting chlorophyll yield, or Chl/TP. Our model explicitly defines the amount of P bound in phytoplankton and suggests that this quantity (Chl/P) varies mainly with eutrophication status. Others studies have observed wide variation in Chl/TP among lakes (Spears et al. 2013). Our model suggests that use of TP to compute this ratio introduces a bias from nonalgal sources of TP that yields generally lower values of Chl/TP and greater variability among lakes. We further found that differences in the P-content of phytoplankton accounted for a small proportion of the overall prediction error, and hence, relative to other contributors to TP, a fixed coefficient reasonably accounts for the contribution of P bound in phytoplankton in most lakes. That is, after accounting for eutrophication status of a lake, the Chl yield of phosphorus (i.e., Chl/P) is nearly constant. The P-content of phytoplankton has been studied extensively and varies with factors such as species composition (Martiny et al. 2013), nutrient availability (Hecky et al. 1993), and light intensity (Sterner et al. 1997). This analysis suggests, however, that in the context of understanding variations in field observations of TP and Chl, factors influencing phytoplankton P-content beyond the species compositional changes occurring with eutrophication are relatively unimportant, which is consistent with the principles of stoichiometric homeostasis (Elser and Sterner 2002). The analysis approach described here may facilitate further comparisons between stoichiometry estimated from lab studies (Klausmeier et al. 2008; Persson et al. 2010) and from field data (Yuan and Jones 2019). Changes in Chl/TP observed in other field studies may also be explained by considering the changes in the proportion of TP associated with phytoplankton. For example, a decrease in Chl/TP has been observed in response to zebra mussel grazing, which is consistent with preferential filtering of phytoplankton from the

water column, leaving inorganic sediment (Nicholls et al. 1999). Changes in Chl/TP observed in Missouri reservoirs in previous analyses are also consistent with an increase in TP associated with SS_{np} (Knowlton and Jones 2000).

Variations in P_{diss} and the P-content of SS_{np} exerted strong effects on the accuracy of model predictions of TP and understanding the causes of these variations would improve predictions in different lakes. Our initial exploration of the effects of lake characteristics on these contributors to TP was not exhaustive but provide the basis for future studies. Most trends were consistent with an understanding of lake processes and with past studies. For example, the strong negative relationship between estimates of P_{diss} and lake depth likely reflects the contribution of resuspended sediment to the overall TP budget near the surface (Krogerus and Ekholm 2003). That is, the likelihood of observing resuspended sediment in the surface layer decreases with greater lake depths (Bloesch 1995). The negative relationship between lake depth and mean SS_{np} (i.e., μ_a) further supports this mechanism. For P_{diss}, this mechanism is predicated on the assumption that a proportion of P_{diss} consists of P bound to sediment fine enough to pass through the filters. Nephelometric measures of filtrate turbidity in previous studies supports this interpretation (Knowlton and Jones 2000).

Model estimates of P_{diss} accurately identified a lower bound to dTP, but in many lakes, dTP was substantially greater than estimated P_{diss}. This difference illustrates the contrast between a statistical approach for estimating components of TP and direct measurements. The direct measurement, dTP, is partly determined by the filter pore size used to extract suspended solids from the whole water sample. Phosphorus passing through the filter is designated as dissolved, whether it is chemically dissolved in the water or bound to fine sediment. Conversely, P_{diss} in the statistical model is defined as any proportion of TP that is uncorrelated with Chl and SS_{np}. In certain lakes, dTP varied directly with SS_{np}, suggesting that the sources of SS_{np} are also the sources of dTP. In these lakes, the statistical estimate of P_{diss} underestimated dTP. In other lakes, luxury uptake of P may have increased the P-content of phytoplankton, but this variability would not be associated with Chl or SS_{np} (Bonachela et al. 2011), and likely would be attributed to P_{diss} in the current model. Neither direct measurement by filtering nor statistical analysis provides an estimate of truly dissolved P or soluble reactive P, nor provide measurements of biologically available P (BAP) or steady-state phosphate concentrations (Butkus et al. 1988; Hudson et al. 2000; Reynolds and Davies 2001). Understanding the biases associated with each method, however, can help interpret their meaning in subsequent analyses.

Differences in filter pore size may slightly influence the applicability of the current results to other data sets, but two factors mitigate this potential effect. First, recent analysis has found that filter pore size does not significantly affect TSS measurements because all pores are clogged similarly by deposited sediment (Kasper et al. 2018). Second, because TP is

modeled as contributions from different compartments, consistency between the filter used to measure Chl and the filter used to measure TSS is a more important consideration than the actual pore size. With our data set, filter pore size was selected to ensure that the same methods could be used in clear lakes in southern MO and turbid lakes in northern MO. Because the pore size used to extract material for Chl analysis was comparable to that used to measure TSS, the data are internally consistent, and estimates of P associated with phytoplankton are robust. The similarity between limiting relationships between Chl and TP derived in MO with those derived in a national data set (collected with finer filters) also supports this conclusion (Yuan and Jones 2019).

Positive relationships between croplands in the watershed with the mean concentration of SS_{np} was consistent with previous observations (Jones et al. 2008a). Surprisingly, though, the amount of P bound to SS_{np} varied considerably among lakes, ranging from 0.1% to 6.2% of the mass of SS_{np}, but no patterns emerged with regard to lake or watershed characteristics. The absence of strong relationships likely suggests that the metrics included in the analysis did not adequately represent the potential causes of differences in sediment P. For example, P bound to allochthonous sediment may be more strongly associated with local farming practices (e.g., timing, amount, and type of fertilizer applied and tillage and cover crops practices), condition of stream banks and severity of stream bank erosion in the watershed, and specific soil characteristics in the watershed (Pote et al. 1996). Also, shoreline erosion can vary substantially depending on fetch and littoral vegetation (Hamilton and Mitchell 1996). The potential for resuspended sediment to contribute to SS_{np} also introduces another source of variability in d_1, as we would expect lakes to sequester different amounts of P in their sediment. The strong effect of variations in d_1 on the accuracy of TP predictions suggests that the collection of some lake-specific data would be necessary to understand changes in TP. Much of the P bound to SS_{np} is not biologically available (Reynolds and Davies 2001), though, and the presence of unknown amounts of P associated with SS_{np} may not be important if the ultimate goal is to predict Chl.

An accurate prediction of Chl resulting from phosphorus loads provides an important tool for managing lake eutrophication, and methods to make these predictions have been the focus of substantial past research. One common approach for making these predictions is to empirically relate P loading directly to Chl concentration (Vollenweider and Kerekes 1982; Rast et al. 1983), such that a targeted Chl concentration can be directly translated into a targeted P load. Others have advocated the use of a two-stage model, in which empirical TP–Chl relationships are first used to estimate TP concentrations that correspond with a desired concentration of Chl (Nicholls 1997). Then, a loading model is used to estimate the P loads that will achieve the desired TP concentration (Ahlgren et al. 1988; Brett and Benjamin 2008). This latter approach may

yield more accurate predictions because an appropriate regional or local TP–Chl relationship could be applied.

Our model suggests that the relationship of Chl yield per unit of P is relatively stable among different lakes, such that an increase or decrease in BAP should result in a predictable change in Chl, regardless of differences in lake characteristics. Other limiting factors (e.g., light, nitrogen) can influence whether all BAP is actually converted to Chl, but the Chl/P relationships specified here can provide the limit imposed by available P. Comparison of Chl predicted by different limiting factors can then potentially provide a simple approach for predicting lake condition (Reynolds and Maberly 2002). Use of this relationship may also improve the accuracy of predictions from yield-based process models (Gowen et al. 1992). Ultimately, the current analysis suggests that the relationship between P and Chl estimated here may be broadly applicable to other locations (Yuan and Jones 2019), and additional analysis of different data sets will further our understanding of factors influencing this relationship.

References

Ahlgren, I., T. Frisk, and L. Kamp-Nielsen. 1988. Empirical and theoretical models of phosphorus loading, retention and concentration vs. lake trophic state, p. 285–303. *In* Phosphorus in Freshwater Ecosystems. Dordrecht: Springer.

Bloesch, J. 1995. Mechanisms, measurement and importance of sediment resuspension in lakes. Mar. Freshw. Res. **46**: 295–304. doi:10.1071/mf9950295

Bonachela, J. A., M. Raghib, and S. A. Levin. 2011. Dynamic model of flexible phytoplankton nutrient uptake. Proc. Natl. Acad. Sci. **108**: 20633–20638. doi:10.1073/pnas.1118012108

Brett, M. T., and M. M. Benjamin. 2008. A review and reassessment of lake phosphorus retention and the nutrient loading concept. Freshw. Biol. **53**: 194–211. doi:10.1111/j.1365-2427.2007.01862.x

Butkus, S. R., E. B. Welch, R. R. Horner, and D. E. Spyridakis. 1988. Lake response modeling using biologically available phosphorus. J. Water Pollut. Control Fed. **60**: 1663–1669.

Carlson, R. E. 1977. A trophic state index for lakes. Limnol. Oceanogr. **22**: 361–369. doi:10.4319/lo.1977.22.2.0361

Deevey, E. S. 1940. Limnological studies in Connecticut; part V, a contribution of regional limnology. Am. J. Sci. **238**: 717–741. doi:10.2475/ajs.238.10.717

Dillon, P. J., and F. H. Rigler. 1974. The phosphorus-chlorophyll relationship in lakes. Limnol. Oceanogr. **19**: 767–773. doi:10.4319/lo.1974.19.5.0767

Effler, S. W., and S. M. O'Donnell. 2010. A long-term record of epilimnetic phosphorus patterns in recovering Onondaga Lake, New York. Fundam. Appl. Limnol. Arch. Für Hydrobiol. **177**: 1–18. doi:10.1127/1863-9135/2010/0177-0001

Elser, J. J., and R. W. Sterner. 2002. Ecological stoichiometry: The biology of elements from molecules to the biosphere.

Princeton University: Press. doi:10.1016/s1074-7613(02)00471-5

Filstrup, C. T., T. Wagner, P. A. Soranno, E. H. Stanley, C. A. Stow, K. E. Webster, and J. A. Downing. 2014. Regional variability among nonlinear chlorophyll—Phosphorus relationships in lakes. Limnol. Oceanogr. **59**: 1691–1703. doi:10.4319/lo.2014.59.5.1691

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge Univ. Press. doi:10.1016/S0072-9752(07)85018-4

Godfrey, P. J. 1982. The eutrophication of Cayuga Lake: A historical analysis of the phytoplankton's response to phosphate detergents. Freshw. Biol. **12**: 149–166. doi:10.1111/j.1365-2427.1982.tb00610.x

Gowen, R., P. Tett, and K. Jones. 1992. Predicting marine eutrophication: The yield of chlorophyll from nitrogen in Scottish coastal waters. Mar. Ecol. Prog. Ser. **85**: 153–161. doi:10.3354/meps085153

Hamilton, D. P., and S. F. Mitchell. 1996. An empirical model for sediment resuspension in shallow lakes. Hydrobiologia **317**: 209–220. doi:10.1007/BF00036471

Havens, K. E., and G. K. Nürnberg. 2004. The phosphorus-chlorophyll relationship in lakes: Potential influences of color and mixing regime. Lake Reserv. Manag. **20**: 188–196. doi:10.1080/07438140409354243

Hecky, R. E., P. Campbell, and L. L. Hendzel. 1993. The stoichiometry of carbon, nitrogen, and phosphorus in particulate matter of lakes and oceans. Limnol. Oceanogr. **38**: 709–724. doi:10.4319/lo.1993.38.4.0709

Hoyer, M. V., and J. R. Jones. 1983. Factors affecting the relation between phosphorus and chlorophyll a in Midwestern reservoirs. Can. J. Fish. Aquat. Sci. **40**: 192–199. doi:10.1139/f83-029

Hudson, J. J., W. D. Taylor, and D. W. Schindler. 2000. Phosphate concentrations in lakes. Nature **406**: 54–56. doi:10.1038/35017531

Jones, J. R., and R. W. Bachmann. 1976. Prediction of phosphorus and chlorophyll levels in lakes. J. Water Pollut. Control Fed. **48**: 2176–2182.

Jones, J. R., and M. F. Knowlton. 2005. Chlorophyll response to nutrients and non-algal seston in Missouri reservoirs and oxbow lakes. Lake Reserv. Manag. **21**: 361–371. doi:10.1080/07438140509354441

Jones, J. R., M. F. Knowlton, and M. S. Kaiser. 1998. Effects of aggregation on chlorophyll-phosphorus relations in Missouri reservoirs. Lake Reserv. Manag. **14**: 1–9. doi:10.1080/07438149809354104

Jones, J. R., M. F. Knowlton, and D. V. Obrecht. 2008*a*. Role of land cover and hydrology in determining nutrients in mid-continent reservoirs: Implications for nutrient criteria and management. Lake Reserv. Manag. **24**: 1–9. doi:10.1080/07438140809354045

Jones, J. R., D. V. Obrecht, B. D. Perkins, M. F. Knowlton, A. P. Thorpe, S. Watanabe, and R. R. Bacon. 2008*b*. Nutrients, seston, and transparency of Missouri reservoirs and oxbow lakes: An analysis of regional limnology. Lake Reserv. Manag. **24**: 155–180. doi:10.1080/07438140809354058

Kasper, D., J. H. Fernandes Amaral, and B. Rider Forsberg. 2018. The effect of filter type and porosity on total suspended sediment determinations. Anal. Methods **10**: 5532–5539. doi:10.1039/C8AY02134A

Klausmeier, C. A., E. Litchman, T. Daufresne, and S. A. Levin. 2008. Phytoplankton stoichiometry. Ecol. Res. **23**: 479–485. doi:10.1007/s11284-008-0470-8

Knowlton, M. F., and J. R. Jones. 1995. Temporal and spatial dynamics of suspended sediment, nutrients, and algal biomass in Mark Twain Lake Missouri. Arch. Hydrobiol. **135**: 145–178.

Knowlton, M. F., and J. R. Jones. 2000. Non-algal seston, light, nutrients and chlorophyll in Missouri reservoirs. Lake Reserv. Manag. **16**: 322–332. doi:10.1080/07438140009354239

Knowlton, M. F., M. V. Hoyer, and J. R. Jones. 1984. Sources of variability in phosphorus and chlorophyll and their effects on use of lake survey data. J. Am. Water Resour. Assoc. **20**: 397–408. doi:10.1111/j.1752-1688.1984.tb04722.x

Krogerus, K., and P. Ekholm. 2003. Phosphorus in settling matter and bottom sediments in lakes loaded by agriculture. Hydrobiologia **492**: 15–28. doi:10.1023/A:1024879309946

Lewis, W. M., and W. A. Wurtsbaugh. 2008. Control of lacustrine phytoplankton by nutrients: Erosion of the phosphorus paradigm. Int. Rev. Hydrobiol. **93**: 446–465. doi:10.1002/iroh.200811065

Martiny, A. C., C. T. A. Pham, F. W. Primeau, J. A. Vrugt, J. K. Moore, S. A. Levin, and M. W. Lomas. 2013. Strong latitudinal patterns in the elemental ratios of marine plankton and organic matter. Nat. Geosci. **6**: 279–283. doi:10.1038/ngeo1757

Mazumder, A. 1994. Phosphorus–chlorophyll relationships under contrasting herbivory and thermal stratification: Predictions and patterns. Can. J. Fish. Aquat. Sci. **51**: 390–400. doi:10.1139/f94-040

McCauley, E., J. A. Downing, and S. Watson. 1989. Sigmoid relationships between nutrients and chlorophyll among lakes. Can. J. Fish. Aquat. Sci. **46**: 1171–1175. doi:10.1139/f89-152

Mellina, E., J. B. Rasmussen, and E. L. Mills. 1995. Impact of zebra mussel (*Dreissena polymorpha*) on phosphorus cycling and chlorophyll in lakes. Can. J. Fish. Aquat. Sci. **52**: 2553–2573. doi:10.1139/f95-246

Molot, L. A., and P. J. Dillon. 1991. Nitrogen/phosphorus ratios and the prediction of chlorophyll in phosphorus-limited lakes in Central Ontario. Can. J. Fish. Aquat. Sci. **48**: 140–145. doi:10.1139/f91-019

Nalewajko, C. 1966. Dry weight, ash, and volume data for some freshwater planktonic algae. J. Fish. Res. Board Can. **23**: 1285–1288. doi:10.1139/f66-119

Nicholls, K. H. 1997. A limnological basis for a Lake Simcoe phosphorus loading objective. Lake Reserv. Manag. **13**: 189–198. doi:10.1080/07438149709354310

Nicholls, K. H., G. J. Hopkins, and S. J. Standke. 1999. Reduced chlorophyll to phosphorus ratios in nearshore Great Lakes waters coincide with the establishment of dreissenid mussels. Can. J. Fish. Aquat. Sci. **56**: 153–161. doi:10.1139/f98-149

Persson, J., P. Fink, A. Goto, J. M. Hood, J. Jonas, and S. Kato. 2010. To be or not to be what you eat: Regulation of stoichiometric homeostasis among autotrophs and heterotrophs. Oikos **119**: 741–751. doi:10.1111/j.1600-0706.2009.18545.x

Phillips, G., O.-P. Pietiläinen, L. Carvalho, A. Solimini, A. Lyche Solheim, and A. Cardoso. 2008. Chlorophyll–nutrient relationships of different lake types using a large European dataset. Aquat. Ecol. **42**: 213–226. doi:10.1007/s10452-008-9180-0

Pote, D. H., T. C. Daniel, P. A. Moore, D. J. Nichols, A. N. Sharpley, and D. R. Edwards. 1996. Relating extractable soil phosphorus to phosphorus losses in runoff. Soil Sci. Soc. Am. J. **60**: 855–859. doi:10.2136/sssaj1996.03615995006000030025x

Prairie, Y. T., C. M. Duarte, and J. Kalff. 1989. Unifying nutrient⊠chlorophyll relationships in lakes. Can. J. Fish. Aquat. Sci. **46**: 1176–1182. doi:10.1139/f89-153

Prepas, E. E., and D. O. Trew. 1983. Evaluation of the phosphorus–chlorophyll relationship for lakes of the Precambrian shield in western Canada. Can. J. Fish. Aquat. Sci. **40**: 27–35. doi:10.1139/f83-005

Pridmore, R. D., W. N. Vant, and J. C. Rutherford. 1985. Chlorophyll-nutrient relationships in North Island lakes (New Zealand). Hydrobiologia **121**: 181–189. doi:10.1007/BF00008722

Quirós, R. 1990. Factors related to variance of residuals in chlorophyll — Total phosphorus regressions in lakes and reservoirs of Argentina. Hydrobiologia **200–201**: 343–355. doi:10.1007/BF02530352

Rast, W., R. A. Jones, and G. F. Lee. 1983. Predictive capability of U.S. OECD phosphorus loading-eutrophication response models. J. Water Pollut. Control Fed. **55**: 990–1003. doi:10.2307/25042007

Reynolds, C. S., and P. S. Davies. 2001. Sources and bioavailability of phosphorus fractions in freshwaters: A British perspective. Biol. Rev. **76**: 27–64. doi:10.1017/S1464793100005625

Reynolds, C. S., and S. C. Maberly. 2002. A simple method for approximating the supportive capacities and metabolic constraints in lakes and reservoirs. Freshw. Biol. **47**: 1183–1188. doi:10.1046/j.1365-2427.2002.00839.x

Sakamoto, M. 1966. Primary production by phytoplankton community in some Japanese lakes and its dependence on lake depth. Arch. Hydrobiol. **62**: 1–28.

Smith, V. H. 1982. The nitrogen and phosphorus dependence of algal biomass in lakes: An empirical and theoretical analysis. Limnol. Oceanogr. **27**: 1101–1112. doi:10.4319/lo.1982.27.6.1101

Smith, V. H., and J. Shapiro. 1981. Chlorophyll-phosphorus relations in individual lakes. Their importance to lake restoration strategies. Environ. Sci. Technol. **15**: 444–451. doi:10.1021/es00086a009

Spears, B. M., L. Carvalho, B. Dudley, and L. May. 2013. Variation in chlorophyll a to total phosphorus ratio across 94 UK and Irish lakes: Implications for lake management. J. Environ. Manage. **115**: 287–294. doi:10.1016/j.jenvman.2012.10.011

Stabel, H.-H. 1986. The role of plankton biomass in controlling fluctuations of suspended matter in Lake Constance. Hydrobiologia **140**: 173–181. doi:10.1007/BF00007572

Stan Development Team. 2016. Stan Modeling Language Users Guide and Reference Manual, Version 2.14.0.

Sterner, R. W., J. J. Elser, E. J. Fee, S. J. Guildford, and T. H. Chrzanowski. 1997. The light: Nutrient ratio in lakes: The balance of energy and materials affects ecosystem structure and process. Am. Nat. **150**: 663–684. doi:10.1086/an.1997.150.issue-6

Vollenweider, R. A., and J. Kerekes. 1982. Eutrophication of waters. Monitoring, Assessment and Control. Paris: OECD.

Wagner, T., P. A. Soranno, K. E. Webster, and K. S. Cheruvelil. 2011. Landscape drivers of regional variation in the relationship between total phosphorus and chlorophyll in lakes. Freshw. Biol. **56**: 1811–1824. doi:10.1111/j.1365-2427.2011.02621.x

Yuan, L. L., and J. R. Jones. 2019. A Bayesian network model for estimating stoichiometric ratios of lake seston components. Inland Waters **9**: 61–72. doi:10.1080/20442041.2019.1582957

Yuan, L. L., and A. I. Pollard. 2014. Classifying lakes to improve precision of nutrient–chlorophyll relationships. Freshw. Sci. **33**: 1184–1194. doi:10.1086/678465

Conflict of Interest
None declared.