Water Research 242 (2023) 120182

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Trophic status and lake depth play important roles in determining the nutrient-chlorophyll *a* relationship: Evidence from thousands of lakes globally

Lei Zhao [a,b,*], Rao Zhu [a], Qichao Zhou [c], Erik Jeppesen [c,d,e,f,g], Kun Yang [a,b,*]

[a] Faculty of Geography, Yunnan Normal University, Kunming 650500, China
[b] GIS Technology Engineering Research Centre for West-China Resources and Environment, Ministry Education, Yunnan Normal University, Kunming 650500, China
[c] Institute for Ecological Research and Pollution Control of Plateau Lakes, School of Ecology and Environmental Sciences, Yunnan University, Kunming 650500, China
[d] Department of Ecoscience, Aarhus University, Aarhus 8000C, Denmark
[e] Sino-Danish Centre for Education and Research, Beijing 100049, China
[f] Department of Biological Sciences and Centre for Ecosystem Research and Implementation, Limnology Laboratory, Middle East Technical University, Ankara 06800, Turkey
[g] Institute of Marine Sciences, Middle East Technical University, Mersin, Turkey

ARTICLE INFO

Keywords:
Eutrophication
Nutrients
Chlorophyll *a*
Lake depth
Bayesian networks
Bayesian hierarchical linear regression model

ABSTRACT

A fundamental problem in lake eutrophication management is that the nutrient-chlorophyll *a* (Chl *a*) relationship shows high variability due to diverse influences of for example lake depth, lake trophic status, and latitude. To accommodate the variability induced by spatial heterogeneity, a reliable and general insight into the nutrient-Chl *a* relationship may be achieved by applying probabilistic methods to analyze data compiled across a broad spatial scale. Here, the roles of two critical factors determining the nutrient-Chl *a* relationship, lake depth and trophic status, were explored by applying Bayesian networks (BNs) and a Bayesian hierarchical linear regression model (BHM) to a compiled global dataset from 2849 lakes and 25083 observations. We categorized the lakes into three groups (shallow, transitional, and deep) according to mean and maximum depth relative to mixing depth. We found that despite a stronger effect of total phosphorus (TP) and total nitrogen (TN) on Chl *a* when combined, TP played a dominant role in determining Chl *a*, regardless of lake depth. However, when the lake was hypereutrophic and/or TP was >40 μg/L, TN had a greater impact on Chl *a*, especially in shallow lakes. The response curve of Chl *a* to TP and TN varied with lake depth, with deep lakes having the lowest yield Chl *a* per unit of nutrient, followed by transitional lakes, while shallow lakes had the highest ratio. Moreover, we found a decrease of TN/TP with increasing Chl *a* concentrations and lake depth (represented as mixing depth/mean depth). Our established BHM may help estimating lake type and/or lake-specific acceptable TN and TP concentrations that comply with target Chl *a* concentrations with higher certainty than can be obtained when bulking all lake types.

1. Introduction

Freshwater eutrophication is a major global problem, and phytoplankton blooms in lakes have increased globally since the 1980s (Ho et al., 2019). Enrichment of nutrients, including phosphorus (P) and nitrogen (N), is considered the main driver of phytoplankton growth. Based on the strong empirical relationship between phytoplankton biomass and chlorophyll *a* (Chl *a*) and P, P reduction with the aim of decreasing the phytoplankton biomass is often the first measure taken to control eutrophication; this is a long-standing paradigm in lake management (Carpenter, 2008; Filstrup et al., 2014). However, several studies have challenged this P control paradigm, pointing out that the effect of N on lake ecosystems is underestimated. A large-scale meta-analysis of experimental enrichment showed that N and P limitation is equivalent in almost all situations (Elser et al., 2007). Studies based on hundreds of lakes in 17 north-eastern and mid-west US states found that

---

Abbreviations: EU, Eutrophication; Nut, Nutrients; Chl *a*, Chlorophyll *a*; LD, Lake depth; BNs, Bayesian networks; BHM, Bayesian hierarchical linear regression model.

* Corresponding authors at: Faculty of Geography, Yunnan Normal University, Kunming 650500, China.
E-mail addresses: zhaolei@ynnu.edu.cn (L. Zhao), yangkun@ynnu.edu.cn (K. Yang).

https://doi.org/10.1016/j.watres.2023.120182
Received 17 March 2023; Received in revised form 20 May 2023; Accepted 6 June 2023

Table 1

Number of lakes, number of observations, mean concentrations (± standard deviation) of observed chlorophyll *a* (Chl *a*), total phosphorus (TP), and total nitrogen (TN) in shallow, transitional, and deep lakes with oligo-mesotrophic, eutrophic, and hypereutrophic status. Lake trophic status was determined according to the threshold values of Chl *a* (oligo-mesotrophic: Chl $a \leq 7$ μg/L, eutrophic: $7 <$ Chl $a \leq 30$ μg/L, hypereutrophic: Chl $a > 30$ μg/L) proposed by USEPA (2009). Shallow lakes were those where the mixing depth exceeded the maximum depth, the transitional lakes were those where mean depth ≤ mixing depth ≤ maximum depth, and the deep lakes were those where the mixing depth was lower than the mean depth.

	Trophic status	Number of lakes	Number of observed data	Mean TN (μg/L)	Mean TP (μg/L)	Mean Chl a (μg/L)
Shallow lakes	Oligo-mesotrophic lakes	77	527	778(±538)	33.2 (±37.2)	4.6 (±1.6)
Shallow lakes	Eutrophic lakes	126	975	1130 (±584)	81.5 (±134.6)	15.9 (±6.8)
Shallow lakes	Hypereutrophic lakes	147	985	2177 (±1108)	162.3 (±125.4)	76.4 (±47.5)
Transitional lakes	Oligo-mesotrophic lakes	496	4568	556 (±428)	17.6 (±18.0)	3.9 (±1.5)
Transitional lakes	Eutrophic lakes	414	3902	1046 (±997)	44.1 (±35.8)	15.2 (±6.5)
Transitional lakes	Hypereutrophic lakes	262	1861	2092 (±1284)	133.7 (±122.7)	63.8 (±39.1)
Deep lakes	Oligo-mesotrophic lakes	734	6530	561 (±511)	16.1 (±15.1)	3.5 (±1.5)
Deep lakes	Eutrophic lakes	448	3933	1163 (±981)	47.6 (±51.6)	14.9 (±6.6)
Deep lakes	Hypereutrophic lakes	145	1802	1799 (±1280)	96.4 (±76.5)	54.3 (±29.9)

P limitation is more likely to occur under oligo-mesotrophic or eutrophic Chl *a* conditions, while co-limitation of P and N occurs at hypereutrophic Chl *a* (Liang et al., 2020). Many case studies have also found successful eutrophication control by combined N and P reduction or reduction of N alone (Abell et al., 2010; Kolzau et al., 2014; Maberly et al., 2020). As different methods and strategies are needed for N and P control and the associated costs differ, it is critical to select an appropriate strategy for P, N or N+P control in lake eutrophication management.

Lake morphological characteristics, such as lake depth may also play a critical role in determining nutrient-Chl *a* relationship and lake eutrophication migration strategy. Søndergaard et al. (2017) and Qin et al. (2020) pointed out that differences in in-lake nutrient dynamics between shallow and deep lakes may play a critical role in determining the response of Chl *a* to nutrients. Internal P loading, for example, is especially important in shallow lakes (Søndergaard et al., 2003). There is also evidence that the role of N might be stronger in shallow lakes as the loss by denitrification is higher here due to e.g. better contact between water and sediment, making N a potential limiting factor in summer (Jeppesen et al., 2007; Søndergaard et al., 2017). Moreover, there is a higher risk of loss of submerged macrophytes at high N as evidenced in several experiments (e.g. Olsen et al., 2015; Søndergaard et al., 2007), being supported also by field data (Jeppesen et al., 2007). Furthermore, Phillips et al. (2008) revealed that shallow lakes have a higher yield of Chl *a* per nutrient unit of P and N than deep lakes, implying that lower nutrient levels are needed in shallow lakes than in deep lakes to pass a desired Chl *a* concentration. Therefore, setting appropriate nutrient criteria based on the response of Chl *a* to nutrients for different lake types is a fundamental task for lake managers to achieve cost-effective eutrophication control.

Many studies have used historical observations, whole-lake experiments in a single lake, or observations in a few selected lakes to identify the limiting factors for phytoplankton growth and to investigate the potential control strategies and responses of Chl *a* to nutrients (Paerl et al., 2016; Qin et al., 2019; Schindler et al., 2016). While providing useful information, inferences from small-scale geographical observations may be constrained to local ecological conditions (Liang et al., 2020) and limited to individual lakes or sets of lakes (Phillips et al., 2008) as the nutrient-Chl *a* relationship may vary temporally and geographically (Abell et al., 2010; Paerl et al., 2016; Canfield et al., 2019; Jeppesen et al., 2020). More universal relationship can be deduced by using datasets of lakes covering large ecological and climate gradients (Janssen et al., 2019, Jeppesen et al., 2020). Some studies based on local scale datasets have been conducted using data-driven predicted models, such as regression equations to investigate the nutrient-Chl *a* relationship and to determine the nutrient that is limiting the phytoplankton growth (Abell et al., 2010; Elser et al., 2007; Huang et al., 2020; Liang et al., 2020; Phillips et al., 2008; Poikane et al., 2019). However, a comprehensive analytical framework suitable for handling the uncertainty induced by spatial heterogeneity when exploring the relationship between nutrients and Chl *a* is lacking.

A model based on data covering broad conditions and a comprehensive analytical framework may provide more general and reliable insight into the nutrient-Chl *a* relationship and, moreover, be directly applied to different type of lakes, including lakes with limited data or low-frequency sampling. To elucidate the complex processes associated with the nutrient-Chl *a* relationship as well as to accommodate the variability of nutrient-Chl *a* induced by spatial heterogeneity, we used a probabilistic machine-learning analytical framework, Bayesian networks (BNs) (Pearl, 1985), and a Bayesian hierarchical linear regression model (BHM) (Malakoff, 1999), an approach widely used in ecosystems modelling (e.g., Cha et al. 2016, Feki-Sahnoun et al. 2017, Malve and Qian 2006, Mellios et al. 2020, Liang et al. 2020). The Bayesian hierarchical linear regression model (BHM) estimate parameters by integrating information from multiple sources and handle uncertainty in a probabilistic manner (Chen and Pollino, 2012; Malve and Qian, 2006; Mellios et al., 2020). Moreover, methods based on Bayes' theorem provide a mutually verified analytical framework that ensures reliable results. Thus, the objectives of this study were to: (1) elucidate the role of N, P, or both are the potential most important growth limiting factor for phytoplankton (Chl *a*) in lakes across broad spatial scales and trophic status gradients; (2) examine whether the potential limiting nutrient varies with lake depth; (3) test BHM to estimate how changes in nutrient concentrations may alter the probability of Chl *a* concentrations to provide detailed information for lake eutrophication management decisions and nutrient criteria; and (4) examine to what extent the response curve of Chl *a* to nutrient concentrations varies with lake depth. The analysis is based on a dataset comprising 2849 lakes globally.

2. Materials and methods

2.1. Dataset

We integrated several regional lake monitoring datasets, including the geospatial and temporal database of US lakes (LAGOS) (Soranno et al., 2017), European lake monitoring data collected by the European Environment Agency member countries (Waterbase-Lakes, https://www.eea.europa.eu/data-and-maps/data/waterbase-lakes-10), the China Lake Scientific Database (http://www.lakesci.csdb.cn), and water quality data from Yunnan Plateau lakes monitored by the Environmental Monitoring Station of Yunnan Province. We conducted a quality control of the compiled data as follows: First, we removed extreme values, defined as Chl $a > 400$ μg/L and < 1 μg/L. We then selected data collected from June to September (the growing season) to minimize the influence of temperature discrepancy caused by diverse latitudes and altitudes. Moreover, to investigate potential differences in the nutrient-Chl *a*

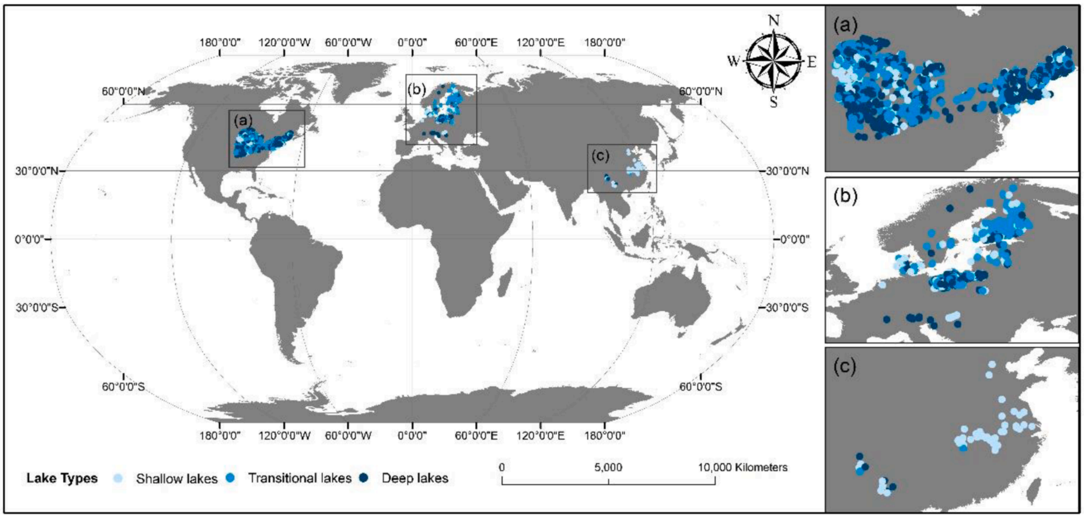

Fig. 1. The spatial distribution of the lakes in our dataset. The lakes were categorized into three groups (shallow, transitional, and deep) according to mean depth and maximum depth relative to mixing depth. Shallow lakes were those where the mixing depth exceeded the maximum depth, the transitional lakes were those where mean depth ≤ mixing depth ≤ maximum depth, and the deep lakes were those where the mixing depth was lower than the mean depth.

relationship related to lake morphometry, recordings of surface area and maximum and mean depth of a number of lakes without documented data on these parameters were drawn from the HydroLAKES database (Messager et al., 2016) according to their geographical location. A total of 25083 observations of TN, TP, and Chl *a* from 2849 lakes (Table 1, Fig. 1) were used in the analysis. Of these 2849 lakes, 1324 had only one observation, while for the remaining lakes the observations ranged from 2 to 354, the average number per lake being 8.8. There are 21 lakes that have more than 100 observations. In our database, 2088 lakes were from the US, 705 from Europe, and 56 from China.

2.2. Categorization of lakes

The categorization of lakes for lake-specific models is often based on the assumption that lakes belonging to a particular type exhibit a similar response of Chl *a* to the stressor (TP and/or TN) (Malve and Qian, 2006; Canfield et al., 2019; Mellios et al., 2020). For example, due to limited information on geographical and natural characteristics of their study lakes, Malve and Qian (2006) classified 2289 Finnish lakes into nine groups based on mean depth and humic type, and Qin et al. (2020) classified their study lakes into three groups based on lake depth. However, for more global datasets, detailed information of lake characteristics are lacking. It is critical to apply effective tools to quantify the Chl *a* nutrient response to the most influential factors, while taking into account the data limitation that often exists when exploring these nutrient-Chl *a* relationships at macroscale.

As numerous local studies have shown that lake depth plays an important role in the Chl *a* nutrient response (Malve and Qian, 2006; Mellios et al., 2020; Søndergaard et al., 2017; Qin et al., 2020), we divided our lakes according to depth following Qin et al. (2020) into three groups: shallow, transitional, and deep based on lake mixing depth (*Zmix*), mean depth (*Zmean*), and maximum depth (*Zmax*), i.e. shallow lakes are those with mixing depths exceeding the maximum depth, transitional lakes are those with a mean depth ≤ mixing depth ≤ maximum depth, and the deep lakes are those with mixing depths lower than the mean depth. The mixing depth was estimated using lake area, $Zmix = (10^{0.185\mathrm{Log}(A) + 0.842} - 2.37) / 1.05$ (where A is the area of the lake surface (km^2)) (Hanna, 1990).

2.3. Bayesian networks

Being a widely used and powerful probabilistic graphical modeling technique, the Bayesian networks (BNs) method (Pearl, 1985) allows investigation of the implicit relationship between variables in a probabilistic manner even when the relationship involves uncertainty. The method is particularly useful when data are integrated across varied ecological conditions. Therefore, a BNs model was established to investigate whether the nutrient limitation of Chl *a* varied with lake type.

A BNs model uses nodes (variables of interest, including parent nodes and child nodes) and edges (the connections among nodes) to represent dependency relationships among nodes. We developed a simple BNs model with TN, TP, and lake depth as drivers (parent nodes) and Chl *a* as the response variable (child node). Although continuous variables can be dealt with to some extent in BNs (Uusitalo, 2007), the usual solution is to discretize the variables to avoid overly complicated algorithms and time-consuming computation (Aguilera et al., 2012; Feki-Sahnoun et al., 2017; (Liang et al., 2020; Uusitalo, 2007). We adapted a lake eutrophication-management-oriented method for discretization of TN, TP, and Chl *a* which defined TN, TP, and Chl *a* as oligo-mesotrophic (TP ≤ 25 μg/L, TN ≤ 750 μg/L, Chl *a* ≤ 7 μg/L), eutrophic (25 < TP ≤ 50 μg/L, 750 < TN ≤ 1400 μg/L, 7 < Chl *a* ≤ 30 μg/L) and hypereutrophic (TP > 50 μg/L, TN > 1400 μg/L, Chl *a* > 30 μg/L) (USEPA, 2009). The discretization of the lake depth was the same as that of lake classification. The discretization data on TN, TP, Chl *a*, and lake depth (TN, TP, and Chl *a* are expressed as trophic status and lake depth as water depth type) were used as input and output for the BNs model. Thus, the established BNs model has three network parent nodes (TN, TP and lake depth) with three states of each parent node and one network child node (Chl *a*) with three states.

We used the R package of bnlearn to estimate the parameters and evaluate model performance (Scutari, 2010). A 10-fold cross-validation strategy, in which the model was fitted 10 times to 90% observations, while the remaining 10% was retained for out-of-sample prediction, was adapted to assess the established model classification accuracy. The classification accuracy was 68.8% which is satisfactory and ensures the reliability of established BNs model. The calibrated BNs model can provide the possibility of Chl *a* trophic status under certain trophic status of TN and TP, and certain lake depth type. For example, we can get the probability of Chl *a* being eutrophic when setting TN to be eutrophic, TP to be hypereutrophic and lake depth to be shallow.

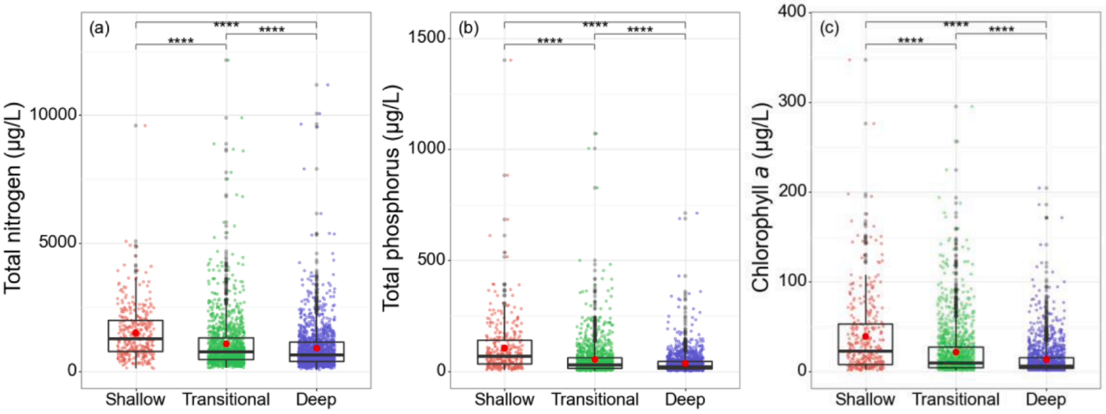

Fig. 2. Boxplots and scatterplots depicting the variability and distribution of (a) total phosphorus (TP), (b) total nitrogen (TN), and (c) chlorophyll *a* (Chl *a*) concentrations for three lake groups (shallow, transitional, and deep). Black line within the box represents the median, red dot within the box represents the mean, black box encompasses the 25th–75th percentiles, and the black dots outside the box represents outliers. Four asterisks indicate the difference between groups are of statistical significance (*Permutation*-test, $p < 0.0001$). Shallow = shallow lakes, Transitional = transitional lakes, Deep = deep lakes. See Fig. 1 for explanations of the categorization of lakes according lake depth.

2.4. Bayesian hierarchical linear regression model

The Bayesian hierarchical linear regression model (BHM) has proven effective and suitable in studies modeling the nutrient-Chl *a* relationship (Malve and Qian, 2006), a hierarchical model can obtain enough parameters from information pooling based on different levels to form a realistic and hierarchical model without overfitting the data (Gelman et al., 2013). Here we used an established BHM based on a dataset covering different lake depths and trophic states to pool information within and among lake types. The BHM used induced a lake-specific curve of the nutrient-Chl *a* response:

$$\log(\text{Chl } a_{ijk}) \sim Normal\left(\beta_{0,ij} + \beta_{1,ij} \times \log(\text{TP}_{ijk}) + \beta_{2,ij} \times \log(\text{TN}_{ijk}), \tau^2\right) \quad (1)$$

$$\beta_{ij} \sim N(\beta_i, \sigma_i^2) \quad (2)$$

$$\beta_i \sim N(\beta, \sigma^2) \quad (3)$$

$$\beta \sim N(0,\ 10000) \quad (4)$$

$$\sigma_i, \sigma, \tau \sim \text{unif}(0,\ 100) \quad (5)$$

where Chl a_{ijk}, TP_{ijk}, and TN_{ijk} are the *k*th observed Chl *a*, TP, and TN

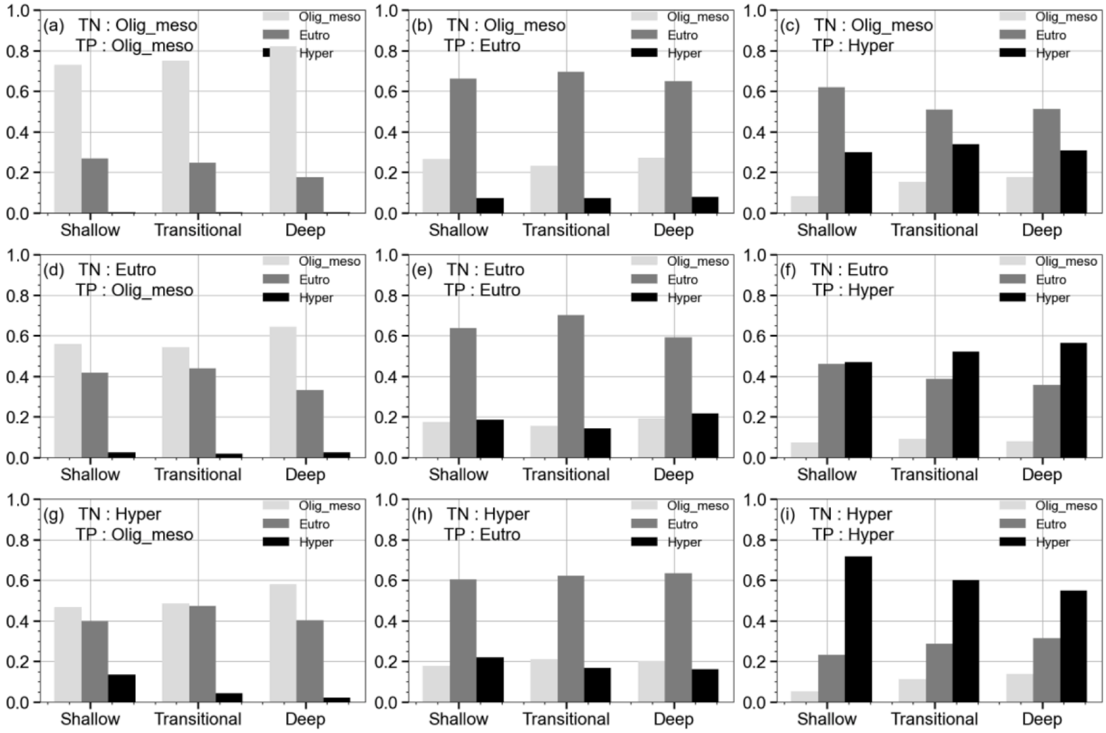

Fig. 3. Probability of chlorophyll *a* trophic status at different combinations of total phosphorus (TP) and total nitrogen (TN) trophic status for three lake groups (shallow, transitional, and deep). These results were inferenced form the established Bayesian networks (BNs) model which was validated using 10-fold cross-validation method and the classification accuracy was 68.8%. Olig_meso = oligo-mesotrophic, Eutro = eutrophic, Hyper = hypereutrophic, Shallow = shallow lakes, Transitional = transitional lakes, Deep = deep lakes. See Fig. 1 for explanations of the categorization of lakes according lake depth.

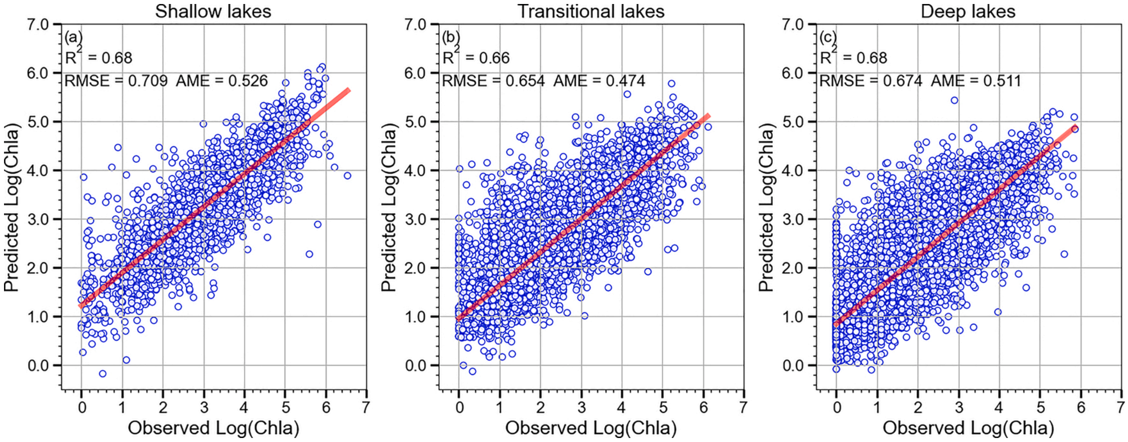

Fig. 4. Model fits (R squared (R^2), Root mean square error (RMSE), and Absolute mean error (AME)) of the Bayesian hierarchical linear regression model (BHM) for three lake groups: (a) shallow lakes, (b) transitional lakes, and (c) deep lakes. The blue scatter plots represent the observed log transformation of chlorophyll *a* (Observed Log(Chla)) vs. the predicted log transformation of the 50th percentile chlorophyll *a* (Predicted Log(Chla)). See Fig. 1 for explanations of the categorization of lakes according lake depth.

concentrations, which are monitored in the *j*th lake ($j = 1, \ldots, 2849$) of type *i* (i = shallow, transitional, deep); τ^2 is the model error variance; β_0 is the intercept; β_1 and β_2 are model slope parameters; βij is the lake-specific model parameter vector consisting of intercept and slopes as well as the parameter for individual lakes at the bottom of the hierarchy; σ_i is the variance of model parameters among lakes belonging to lake type i; σ is the variance among lake types; β is the global parameter with a normal distribution having a mean of 0 and a variance of 10000; unif (0, 100) is the uniform distribution of σ_i, σ, and τ.

Parameter estimation and model implementation were based on the python library of PyMC3, which was developed for Bayesian statistical modeling and probabilistic machine learning (Salvatier et al., 2016). Coefficient of determination (R squared (R^2)) , absolute mean error (AME) and root mean square error (RMSE) were used to assess the accuracy and precision of the model fit.

3. Results

3.1. The role of TP and TN in differential depth lakes

The statistics and boxplots of TN, TP, and Chl *a* for the different groups of lakes are shown in Table 1 and Fig. 2. We observed an increase in the growing season concentrations of TN, TP, and Chl *a* with decreasing lake depth, and the difference between shallow, transitional and deep lakes were statistically significant (*Permutation*-test, $p < 0.0001$). Even when a lake had the same trophic status, the concentrations of TN, TP, and Chl *a* in the shallow lakes were highest among the three lake groups. Although a higher mean Chl *a* concentration was related to higher mean TN and TP concentrations in all groups, lake depth seems to play an important role for the algal biomass.

We explored the role of lake depth for the nutrient-Chl *a* relationship using the results of established BNs model for the probabilities of Chl *a* trophic status under different combinations of TN and TP trophic states (Fig. 3). We found that 27.0% of the shallow lakes were classified as eutrophic or hypereutrophic when the nutrient state was oligo-mesotrophic, while 25.1% of the transitional lakes and 17.9% of the deep lakes were in belonged to these categories (Fig. 3(a)). Moreover, when the trophic status of TN and TP changes from oligo-mesotrophic to eutrophic or hypereutrophic, the probability of Chl *a* trophic status being eutrophic or hypereutrophic increased sharply in all lake groups. Among the three groups, 82.6% of the shallow lakes, 84.4% of the transitional lakes, and 80.7% of the deep lakes were now eutrophic or hypereutrophic state when the nutrients reached the eutrophic state, while 95.0% of the shallow lakes, 88.7% of the transitional lakes, and 86.2% of the deep lakes were eutrophic or hypereutrophic when the nutrients reached a hypereutrophic state (Fig. 3(a), 3(e), and 3(i)). These results showed that the increase in TN and TP concentrations in lakes induces an increase in the Chl *a* concentration at macroscale, being most pronounced in shallow lakes.

3.2. The role of TP and TN in differential trophic status lakes

Comparing Chl *a* trophic status with one nutrient fixed and one variable allowed us to explore the relative effect of nutrients on Chl *a*. As shown in Fig. 3(a), 3(b), and 3(c), an oligo-mesotrophic TN trophic state that changes the TP state from oligo-mesotrophic to eutrophic or hypereutrophic resulted in a decrease in the probability of Chl *a* being oligo-mesotrophic for all lakes (probability of a decline 0.65, 0.60, and 0.64 for shallow, transitional, and deep lakes, respectively, when TP is eutrophic) and an increase in the probability of Chl *a* tropic status being eutrophic and hypereutrophic. In contrast, when the constant TP trophic state was oligo-mesotrophic, and the TN state changed from oligo-mesotrophic to eutrophic or hypereutrophic, there was a relatively small decrease in the probability of Chl *a* trophic status being oligo-mesotrophic. For instance, a shift in TN from oligo-mesotrophic to hypereutrophic led to a 26.3% (0.73 – 0.47) shift of shallow lakes from an oligo-mesotrophic to a eutrophic or hypereutrophic state, which is smaller than the transfers resulting from a shift in TP (64.6%, 0.73 – 0.08). These results show that while both TN and TP influence Chl *a*, TP seems to overall play a more important role. Consistent with the above-mentioned results, when TP and TN trophic statuses are hypereutrophic and oligo-mesotrophic, respectively, the probability of Chl *a* being eutrophic or hypereutrophic was as high as 0.92 (shallow lakes), 0.85 (transitional lakes), and 0.82 (deep lakes) (Fig. 3(c)). However, when the TP trophic status was oligo-mesotrophic and the TN status was hypereutrophic, the probability of Chl *a* trophic status being eutrophic or hypereutrophic was relatively low (Fig. 3(g)). Changing the TP trophic status from oligo-mesotrophic to eutrophic or hypereutrophic led to an increase in Chl *a* trophic status to mainly eutrophic or hypereutrophic (Fig. 3(g), 3(h), and 3(i)). These results demonstrate that TP generally plays a more important role than TN in determining Chl *a* trophic status, both for shallow lakes and for deep lakes. However, TN may play an important role in some situations. As displayed in Fig. 3(a) and 3(c), with the change in TN trophic status from oligo-mesotrophic to hypereutrophic when the TP status is oligo-mesotrophic, the trophic status of Chl *a* being hypereutrophic in shallow lakes increases more than in transitional and deep lakes. Similarly, changing TN trophic status from oligo-mesotrophic to hypereutrophic when the TP status is

Table 2
The 50th percentile posterior distribution parameters of the established Bayesian hierarchical linear regression model for the three lake groups. See Table 1 for explanations of the categorization of lakes according lake depth.

Lake Group	β_0	β_1	β_2
Shallow lakes	-0.131	0.591	0.084
Transitional lakes	-0.249	0.574	0.087
Deep lakes	-0.470	0.511	0.117

hypereutrophic led to an increase in Chl *a* trophic status as hypereutrophic in all lakes. These results demonstrate that TN may have a more significant interactive impact on the Chl *a* status in shallow lakes when the TP trophic status is hypereutrophic.

3.3. *Response of Chl a to nutrients based on posterior predictive simulation of BHM*

Having explored the potential effects of TP and TN on Chl *a* with variable lake depth based on the BNs model, we used the BHM to quantify the Chl *a* nutrient response. Taking advantage of the simplicity and flexibility of BHM is useful for eutrophication management (Malve and Qian, 2006; Mellios et al., 2020). To evaluate the model accuracy, we plotted the log transformation of observed versus predicted Chl *a* in the three lake groups. As shown in Fig. 4, the BHM performed rather well for all lake groups; the R^2, RMSE, and AME were 0.68, 0.71, and 0.53 for shallow lakes; 0.66, 0.65, and 0.47 for transitional lakes; and 0.68, 0.67, and 0.51 for deep lakes, respectively. The accuracy of the established BHM provided reliable results and corresponding posterior prediction. Table 2 lists the 50th percentile posterior distribution parameters of BHM for the three lake groups. The coefficient of TN (β_2) is smaller than the coefficient of TP (β_1) in all lake depth groups, which enforces that TP is the more influential than TN on Chl *a*.

Given the possible collinearity of TP and TN, it is difficult to determine their effects on Chl *a* only on the coefficients estimated, the effect of TP or TN on the nutrient-Chl *a* response curve was further investigated by using one nutrient that varied within the 5th to 95th percentile range of the observed data (dataset including all lake groups) and another nutrient that remained constant (25th percentile, threshold for oligo-mesotrophic, threshold for eutrophic, and 95th percentile). To ensure an even distribution of the data interval, the varied nutrient range was divided into four (5th percentile–25th percentile, 25th percentile–50th percentile, 50th percentile–75th percentile, 75th percentile–95th percentile), and each range was discretized into 15. The results of the posterior predictive simulation of the effect of TP or TN on the nutrient-Chl *a* relationship are illustrated in Fig. 5. Overall, the predicted Chl *a* increased with TP and TN, demonstrating that both nutrients can influence Chl *a*. However, increased TP produced a significantly greater increase in Chl *a* than TN, clearly indicating that TP plays a dominant role in determining the Chl *a* concentration across all lake groups and implying that algal growth is more likely limited by TP than by TN. Moreover, although TN influenced Chl *a* less than TP, the effect of TN on Chl *a* increased with increasing TP. (Fig. 5(d), 5(e), and 5(f)).

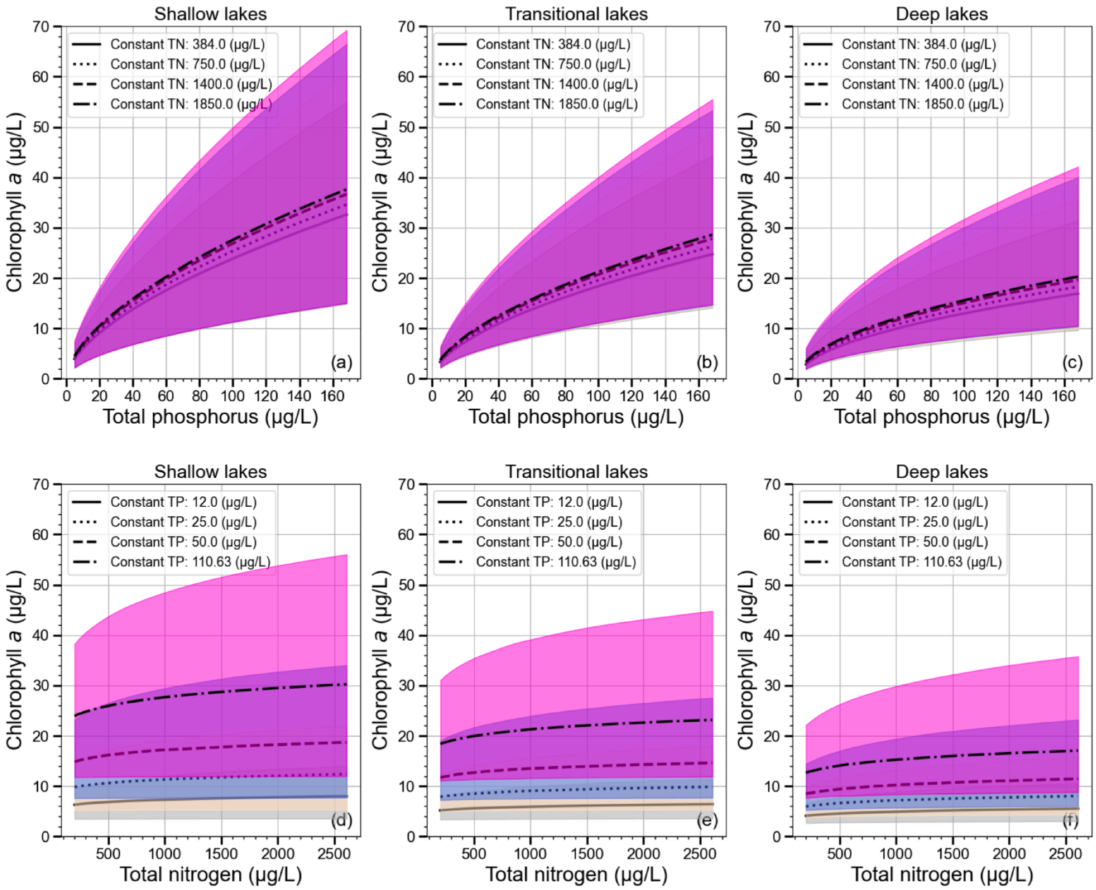

Fig. 5. Chlorophyll *a* concentration as a function of total phosphorus (TP) and total nitrogen (TN) concentrations for shallow ((a) and (d)), transitional ((b) and (e)), and deep ((c) and (f)) lakes, predicted with the lake-type Bayesian hierarchical linear regression model. The colored regions represent the 10 – 90th percentile of the predicted chlorophyll *a* concentration, while the solid and dotted black lines represent the 50th percentile of the predicted chlorophyll *a* concentration. While one nutrient varied within the range of the 5th to 95th percentile of the observed data (dataset included all lake groups), another nutrient remained constant (25th percentile of the observed data, threshold for oligo-mesotrophic, threshold for eutrophic, and 95th percentile of the observed data). Constant TP = constant concentration of total phosphorus; Constant TN = constant concentration of total nitrogen. See Fig. 1 for explanations of the categorization of lakes according lake depth.

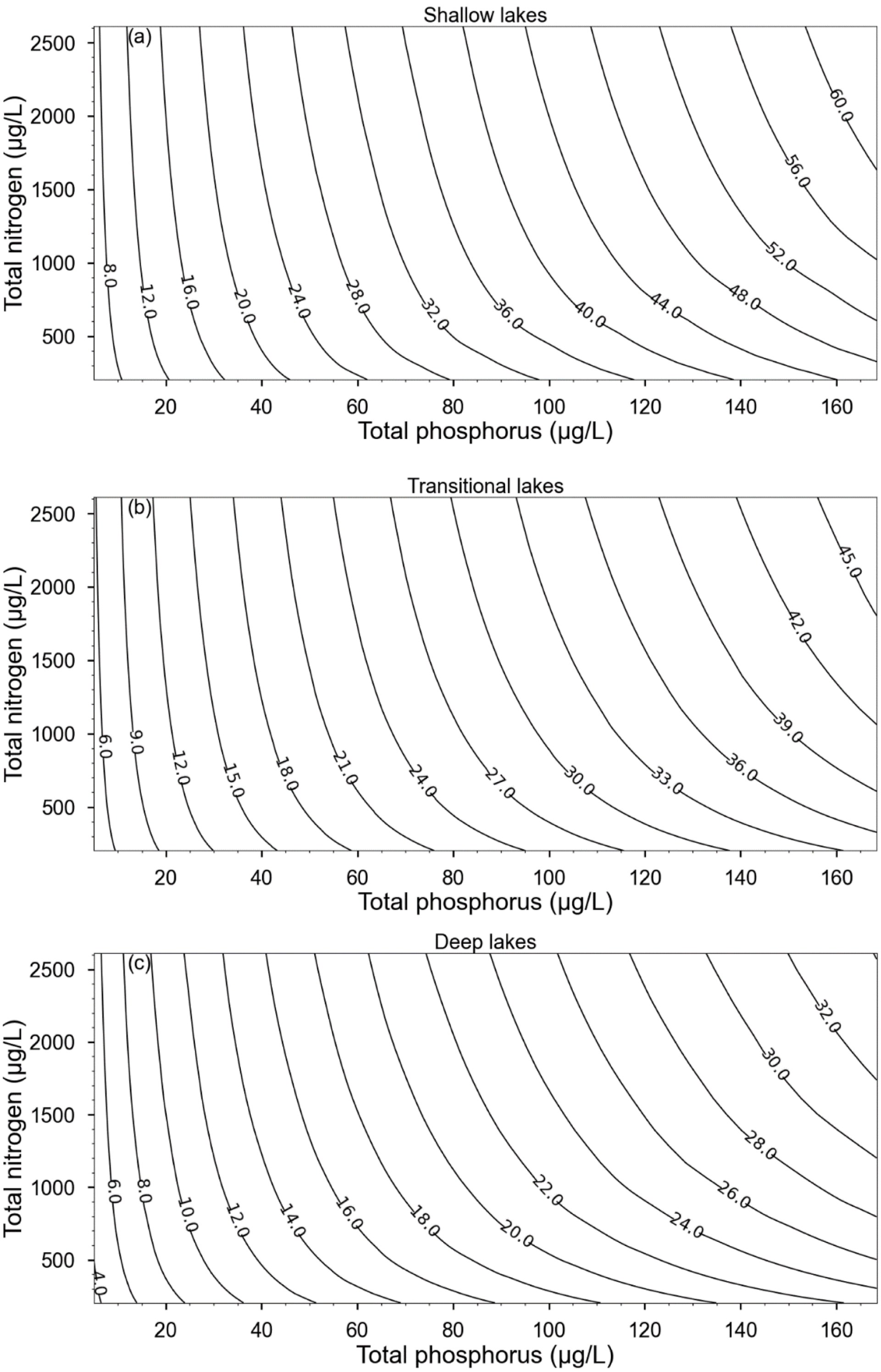

Fig. 6. 50th percentile contour lines of the chlorophyll *a* concentration as a function of total phosphorus (TP) and total nitrogen (TN) concentrations for: (a) shallow lakes, (b) transitional lakes, and (c) deep lakes predicted with the lake-type Bayesian hierarchical linear regression model. TP and TN varied within the range of the 95th and 5th percentile of the observed data (dataset including all lake groups). Numbers are the predicted chlorophyll *a* concentrations (μg/L). See Fig. 1 for explanations of the categorization of lakes according lake depth.

Additionally, at the same level of nutrient concentration, shallow lakes exhibited the highest Chl *a* concentration among all the lake groups. These results show a similar pattern to the inference of the BNs model, lending robustness to our BHM and the credence of the analysis.

To investigate the different nutrient criteria among lake groups, another posterior prediction was implemented based on variations in TN and TP, in which the TN and TP plane within a special range was discretized into 60 × 60 grid cells, and the Chl *a* concentration of each nutrient cell was calculated for all lakes in each lake group. The method for nutrient discretization was the same as in the first implementation. The response curves of the 50th percentile Chl *a* to the TP and TN of the different lake groups were plotted as contour lines (Fig. 6), presenting a clear nutrient-Chl *a* distribution relationship for a specific lake group. Substantial differences in the nutrient-Chl *a* responses were observed within the lake groups and among the groups. First, when TP was <40 μg/L, the contour lines were parallel to the y-axis. As TP increased, the contour lines became more parallel to the x-axis. Moreover, when the concentrations of TP and TN remained the same for different lake groups, the concentration of Chl *a* was the highest in the shallow lakes, followed by transitional lakes and deep lakes.

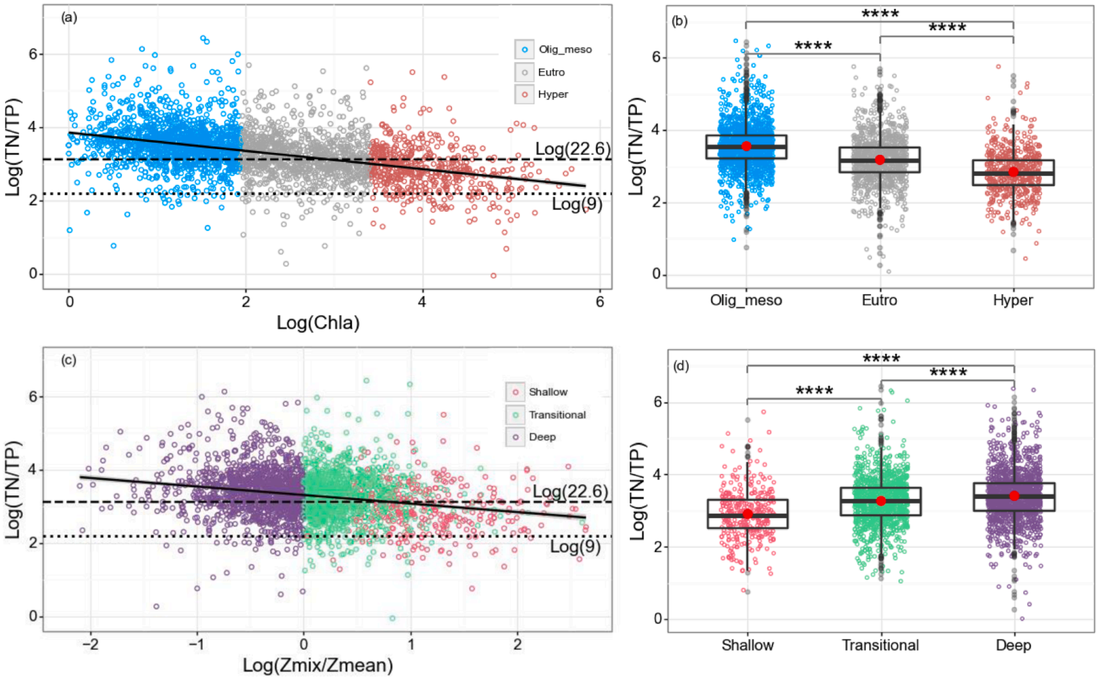

Fig. 7. The relationship between Log(TN/TP) and Log(Chla) (a), Log(Zmix/Zmean) and Log(Chla) (c) for lakes in our dataset. Solid black lines represent the fitted regression line ((a): Pearson correlation coefficient (R) = -0.433, $p < 0.001$; (c): R = -0.228, $p < 0.001$), horizontal dashed lines indicate the threshold of P limitation on Log_e scale, horizontal dotted lines indicate the threshold of N limitation on Log_e scale. Boxplots and scatterplots depicting the variability and distribution of Log (TN/TP) for: (b) three lake trophic status group, (d) three lake depth groups. Lake trophic status was determined according to the threshold values of chlorophyll *a* proposed by USEPA (oligo-mesotrophic: Chl *a* ≤ 7 μg/L, eutrophic: 7 < Chl *a* ≤ 30 μg/L, hypereutrophic: Chl *a* > 30 μg/L) (USEPA, 2009). Olig_meso = oligo-mesotrophic, Eutro = eutrophic, Hyper = hypereutrophic, Shallow = shallow lakes, Transitional = transitional lakes, Deep = deep lakes, Chla = chlorophyll *a*, TN = total nitrogen, TP = total phosphorus, Zmix = mixing depth of lake, Zmean = mean depth of lake. See Fig. 1 for explanations of the categorization of lakes according lake depth. See Fig. 2 for further explanation of the box plots.

We further checked the robustness of the BHM and the above results by randomly selecting nine observations (the average number per lake) from the lakes with observations greater than 300 – Lake 6301 (354 observations) and Lake 6302 (340 observations) in the LAGOS dataset, while the observations of other lakes remained unchanged. We found that the results revealed similar patterns to the BHM trained with the full dataset, as shown in the supplementary material (Figs. S1–S3).

3.4. Nutrient limitation

On the basis of TN/TP, lake tropic status and depth in our dataset, we found an apparent decrease of TN/TP ratio with the increase of Chl *a* concentration (the linear fitted regression in Fig. 7(a), Pearson correlation coefficient (R) = -0.433, $p < 0.001$), and the differences of TN/TP ratio between oligo-mesotrophic, eutrophic and hypereutrophic lakes were statistically significant (*Permutation*-test, $p < 0.0001$). Meanwhile, we found a decrease in the TN/TP ratio with increasing *Zmix/Zmean* (the linear fitted regression in Fig. 7(c), R = -0.228, $p < 0.001$), and the differences of TN/TP ratio between shallow, transitional and deep lakes were statistically significant (*Permutation*-test, $p < 0.0001$). Moreover, as the increase of Chl *a*, the limiting nutrient may shift from only P limitation (TN/TP > 22.6, the horizontal dashed lines in Fig.7(a)) to N + P co-limitation ($9 \leq TN/TP < 22.6$) or only N limitation ($TN/TP \leq 9$, the horizontal dotted lines in Fig. 7(a)). Average TN/TP ratio of oligo-mesotrophic, eutrophic and hypereutrophic are 35.2, 24.0 and 17.1, respectively. A shift from only P limitation to N + P co-limitation or only N limitation with decreasing lake depth was also observed (Fig. 7(c)).

4. Discussion

Our analyses, based on a dataset comprising 2849 lakes and 25083 observations analyzed using Bayesian networks and Bayesian hierarchical linear regression models, showed that synergistic effects of TP and TN on Chl *a* occurred in all three lake types studied. TP played a dominant role in determining Chl *a*, while TN particularly contributed when the lake TP concentration was high, and especially in shallow lakes. Lake depth played an important role in determining nutrient sensitivity, and shallow lakes are those most likely to become eutrophicated, followed by transitional lakes, and then deep lakes.

4.1. In-lake dynamics controlling the nutrient-Chl *a* relationship

Several morphological features such as the proportion of the lake area where the sediment is in direct contact with the water, hypolimnion depth (if any hypolimnion), and the proportion of littoral zones suitable for macrophyte growth vary markedly among shallow, transitional, and deep lakes, and this can profoundly affect the nutrient processes (Diehl, 2002; Genkai-Kato and Carpenter, 2005; Søndergaard et al., 2017). Several model exercises and enclosure experiments have shown that lake morphological characteristics, especially lake depth and mixed depth, play critical roles in regulating the nutrient supply from the sediment, denitrification at the water-sediment interface, and phytoplankton growth (Cantin et al., 2011; Diehl, 2002; Diehl et al., 2002; Genkai-Kato and Carpenter, 2005; Longhi and Beisner, 2009), and such differences may explain why we found different nutrient-Chl *a* relationship for shallow, transitional, and deep lakes. In shallow lakes, which have close sediment contact with the water column and are most frequently disturbed by wind, phytoplankton growth may be particularly supported by additional P and N from the sediment (Søndergaard et al., 2003). Different proportion of littoral zones suitable for macrophyte growth may be another factor affecting variable nutrient-Chl *a* relationship between shallow and deep lakes. In general, through several mechanisms, e.g. stabilizing sediments and further decreasing the rate of nutrient cycling, providing refuge for zooplankton and further enhancing the top-down regulation of phytoplankton by zooplankton (Genkai-Kato and Carpenter, 2005; Jeppesen, et al., 2007; Søndergaard et al., 2017), shallow lakes with abundant macrophytes are considered more resilient to the pressure of nutrient loading and keep low phytoplankton biomass. However, the loss of macrophyte induced by external or internal drivers (Phillips et al., 2016; Søndergaard et al., 2017; Zhang et al., 2017), e.g. enrichment of nutrients, severe fluctuation of water level, increase of benthivorous fish biomass, can lead to shallow lakes experiencing more internal loading of P and higher phytoplankton growth compared to deep lakes. In contrast to shallow lakes, deep lakes have less macrophytes, lower contact between sediment and the productive epilimnion layer and higher mixing depth (meaning potentially less light availability) (Diehl, 2002; Diehl et al., 2002; Dubourg et al., 2015; Yang et al., 2023; Zhao et al., 2021), which all may result in lower phytoplankton growth. Similar to other studies (Phillips et al., 2008), our dataset unfortunately did not contain information on presence/absence of macrophytes, but it is reasonable to assume based on data from numerous Danish lakes (Søndergaard et al., 2017), that a shift from a macrophyte to a phytoplankton state with increasing nutrients concentration explain why the shallow lakes showed a more sensitive response of Chl *a* to enrichment of nutrients than the deep lakes in our dataset.

Key processes leading to higher importance of N for phytoplankton growth in eutrophic lakes are denitrification and P release from the sediment and the feedback between high phytoplankton biomass and the two processes (Muller et al., 2021; Qin et al., 2020; Saunders and Kalff, 2001). An increase in the denitrification rate with increasing trophic status has been widely observed (Muller et al., 2021; Wu et al., 2019). As trophic status increases, settled phytoplankton accumulates and decays in the sediment, providing labile organic matter for denitrification and favoring sediment denitrification and thus N removal. Especially in shallow lakes, turbulence may produce oxic conditions in the water-sediment interface and lead to nitrification. This favors the coupled nitrification-denitrification process and thus accelerates the loss of N (Zhu et al., 2021). Moreover, high accumulation of organic matter in the sediment promotes P release from the sediment to the water column, especially in shallow lakes (Søndergaard et al., 2003), increasing the potential for N limitation (Diehl, 2002; Diehl et al., 2002; Kolzau et al., 2014; Qin et al., 2019, 2020).

The role of N in nutrient-rich lakes, not least the shallow ones, is also evident from several previous lake-specific or multi-lake studies. For example, a study conducted in four German lakes revealed that N was more limiting in summer in shallow lakes; the limiting nutrient alternated between P and N during summer in a deep stratified lake (Kolzau et al., 2014). In Lake Erie, USA, which is a well-documented eutrophic deep lake, phytoplankton biomass was co-limited by P and N in August to October in the central basin (Jessica and Hunter, 2007), and in Maumee Bay (located in the southwest end of Lake Erie and the monitoring site having a depth of 2.5 m), the limiting nutrient of phytoplankton biomass shifted from P to N during the summer season (Chaffin et al., 2014). Moreover, a multi-lake study based on observations of 817 Danish lakes revealed that in both deep and shallow lakes, Chl *a* was more strongly related to TP than to TN; however, TN accounted for more of the variability in Chl *a* than TP when TP concentrations were high (Søndergaard et al., 2017). A study of seasonal variation in limiting factors for Chl *a*, biomass of total phytoplankton biomass and cyanobacteria in 503 Danish lakes revealed an overall dominance of P limitation during winter, early spring and late autumn, while the role of N increased in summer, not least in shallow lakes (Jeppesen et al., 2021). Downing and McCauley (1992) summarized published observations from 221 lakes in 14 countries and concluded that N limitation is significantly more frequent in lakes when TP is greater than 30 μg/L. This finding is similar to our result showing that when $TP > 40$ μg/L, the role of TN determining Chl *a* increased. Because our analysis comprised a dataset covering a broad spatial scale using a robust analytical framework, our findings have the potential to provide more general insight into the nutrient-Chl *a* relationship and the role of trophic status and lake depth in determining the relationship than obtained in previous studies. Although lake-specific strategies are most effective in controlling eutrophication, for many lakes no or only scarce data are available, which makes it difficult to establish a proper eutrophication control strategy. In such cases, our developed models may be a helpful tool.

To obtain simple relationships, we did not include water temperature as an explanatory variable, but recent studies suggest that the yield Chl *a* / (TP or TN) might differ among climate regions, partly because the changes in top-down control of phytoplankton are more pronounced in lakes located in colder regions as a result of lower fish predation on zooplankton (Jeppesen et al., 2020, 2021). Further studies, including data from different climate zones (ours are mainly from north temperate lakes), are needed to elucidate whether inclusion of water temperature can reduce the uncertainties of the relationships presented.

TN and TP were used here as the nutrients indicators in determining the relationship between nutrients and phytoplankton due to data availability, which guarantee a dataset covering a great nutrient gradient and a wide range of ecological conditions, and a more general insight. Meanwhile, TN and TP were also the most widely used nutrients indicators in determining nutrients limitation of phytoplankton and for eutrophication control (Jeppesen et al., 2020, 2021; Liang et al., 2020; Maberly et al., 2020; Malve and Qian, 2006; Mellios et al., 2020; Phillips et al., 2008; Qin et al., 2020). However, phytoplankton often contains an important part of TN and TP in the water, making TN and TP not fully independent variables, and some studies, therefore, suggest that dissolved inorganic nitrogen (DIN) would be a better predictor for available nitrogen of phytoplankton growth than total nitrogen, partly because luxury uptake of N is negligible, and DIN : TP seems the best indicator for nutrient limitation (Morris and Lewis, 1988; Ptacnik et al., 2010). A dataset including different forms of nitrogen and phosphorus, such as DIN, dissolved organic nitrogen (DON), dissolved inorganic phosphorus (DIP) and dissolved organic phosphorus (DOP), is needed to investigate whether inorganic nitrogen forms can predict phytoplankton growth better and reduce the uncertainties of the relationships presented. We concentrated on the most critical summer period with typical high phytoplankton biomasses, but future studies should also include other seasons as the results may differ from those based on summer data only (Jeppesen et al., 2021).

4.2. N vs P limitation and implication for eutrophication control

As the responses of lake ecosystem to nutrients reduction include many slow, long-term changes to nutrients fluxes, nutrients that provide long-term controls on phytoplankton biomass can be termed as *ultimate* limiting nutrients (Schindler, 2012), which should be deduced by long-term ecosystem-scale observations and/or macroscale observations. Our results based on datasets across broad spatial scales, trophic status gradient and lake depth showed that P is dominant in determining Chl *a* and the role of N is secondary, but important in eutrophic lakes, especially the shallow ones. Although our findings cannot be generalized to all lakes, it provides an important and insightful prior information for lake eutrophication management. First, solely P reduction is effective in migrating eutrophication especially in deep lakes, but dual N and P would likely be more advantageous and should be a more effective strategy in shallow eutrophic lakes. In shallow lakes, the enrichment of N reduces the resilience of lake ecosystem, e.g. by imposing high risk of loss of submerged macrophytes (Gonzalez Sagrario, et al., 2005; Olsen et al., 2015; Phillips et al., 2016) which, in turn, increases the risk of algae blooms, and this may delay the re-establishment of oligotrophic or mesotrophic status. Second, given the need for different approaches and related costs, local evaluations are recommended to identify the limiting nutrient of phytoplankton growth and to specify appropriate nutrient criteria for target Chl *a* concentration. Finally, our results clearly show that an adaptive strategy for nutrient control should be taken which

Table 3
Mean depth (MED), maximum depth (MAD), surface area (SA), number of observed data, mean concentration (±standard deviation) of observed chlorophyll *a* (Chl a), total phosphorus (TP) and total nitrogen (TN) of three selected lakes predicted by the established Bayesian hierarchical linear regression model.

Lake name	MED (m)	MAD (m)	SA (km^2)	Mean TN (μg/L)	Mean TP (μg/L)	Mean Chl a (μg/L)	Number of observed data
Lake Dianchi	2.9	5.9	298	2331.8 (±446.2)	153.2 (±61.8)	81.3 (±37.0)	25
Lake Erhai	10.2	25.3	249	567.3 (±59.1)	29.7 (±6.7)	11.2 (±4.9)	45
Lake LAGOS 6302	22.9	64	182	629.7 (±377.8)	32.3 (±33.6)	19.2 (±22.7)	341

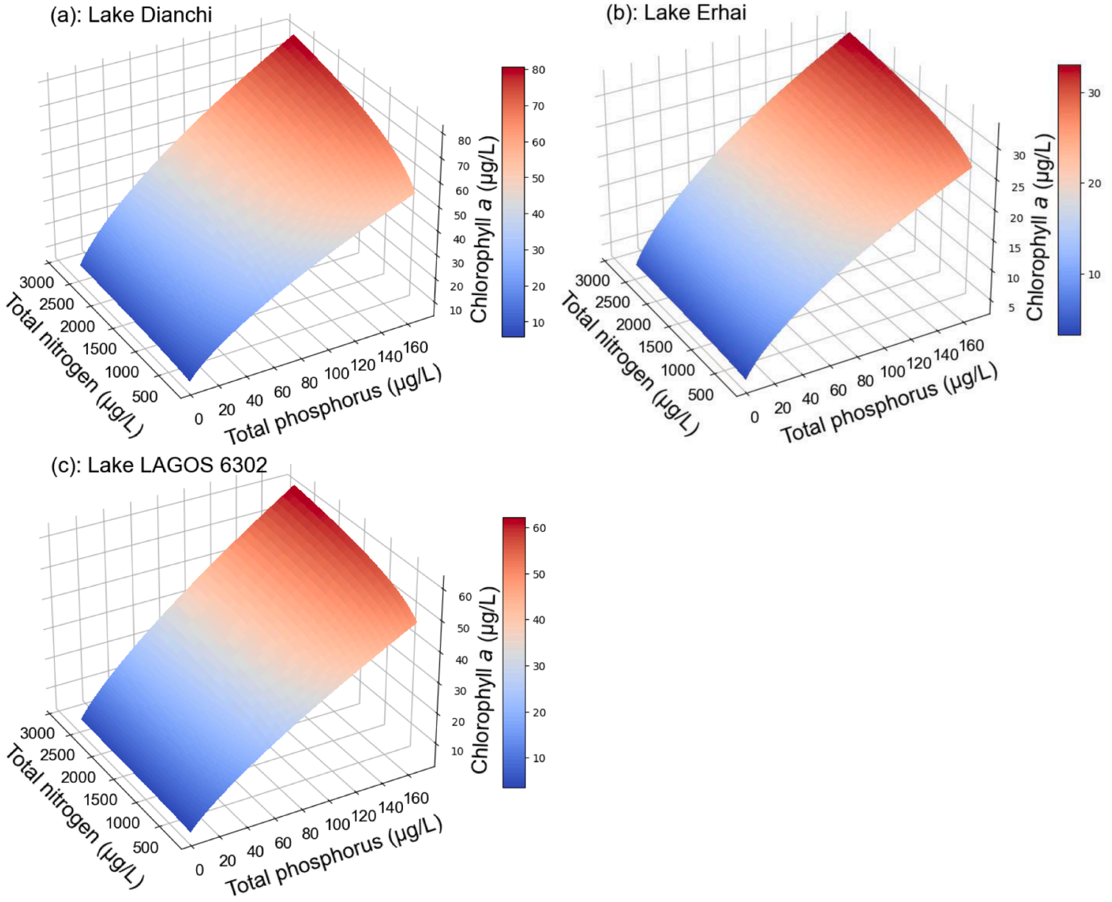

Fig. 8. The 50th percentile response surface of chlorophyll *a* concentrations relative to a combination of total phosphorus (TP) and total nitrogen (TN) concentrations for: (a) Lake Dianchi (shallow), (b) Lake Erhai (transitional), and (c) Lake Lagos 6302 (deep) predicted with the lake-specific Bayesian hierarchical linear regression model. TP and TN varied within the range of the 95th and 5th percentile of the observed data (dataset including all lake groups). Please refer to Table 3 for the morphometric and chemical characteristics of the three lakes.

target nutrient is not fixed but should be identified according to the trophic status and type of the lake.

4.3. Implications of the lake-specific BHM for eutrophication control

As demonstrated by Malve and Qian (2006) and Mellios et al. (2020), BHM can provide information at various hierarchy levels of lakes, e.g. all lakes included, specific lake types and specific lakes depending on data available so as to minimize the uncertainty and bias associated with lake-specific prediction. In addition, the established BHM based on multi-lake observations can function as a powerful tool for eutrophication control by setting lake-specific nutrient criteria that may require driver-response predictions outside the observational range. To illustrate this further, we implemented posterior predictions for three lakes (Lake Dianchi, shallow lake; Lake Erhai, transitional lake; Lake 6302, a lake in the LAGOS dataset (deep lake)) to demonstrate how the BHM can be used to drive lake-specific nutrient criteria. Table 3 lists the morphometric and chemical characteristics of the selected lakes.

The lake-specific BHM showed variable accuracy: R^2, RMSE, and AME were 0.45, 0.35, and 0.27 for Lake Dianchi, 0.31, 0.38, and 0.32 for Lake Erhai, and 0.62, 0.66, and 0.47 for Lake LAGOS 6302. Although the performance of the lake-specific BHM had higher bias than the lake type BHM, they showed high accuracy considering the limited number of observations. We simulated posterior probabilities of Chl *a* for a range of TN and TP values (the range between the 5th and 95th percentiles of all data; TN and TP were discretized into 100 × 100 grids) for these three specific lakes. The surface of the simulated 50th percentile Chl *a* relative to variable TN and TP is plotted in Fig. 8. These surface plots provide clear lake-specific nutrient-Chl *a* response curves, which can be easily understood and used for setting nutrient criteria. Meanwhile, even if the nutrient ranges (TN: 203–2610 μg/L, TP: 5.0–168 μg/L) used for simulation are outside the observational range for the specific lake, reasonable extrapolation can help determine target TP and/or TN values. Moreover, the predicted probabilistic distribution of Chl *a* obtained provides more information than can be obtained by the traditional regression models and are therefore better tools to make decisions within a risk assessment framework. For instance, if we need to balance the financial expenditure and restoration target, the nutrient response of the 50th percentile Chl *a* may be an appropriate choice. If we need to minimize the risk of exceeding a given Chl *a* concentration with no financial considerations, a cautionary decision may be adopted by the decision maker, and the corresponding nutrient target to the 90th percentile Chl *a* can be used.

5. Conclusions

A dataset spanning thousands of lakes across broad spatial scales, trophic status gradient and lake depth was analyzed applying an established Bayesian analytical framework. The dataset was grouped according to lake depth and mixing depth to explore the effects of lake depth and trophic status on the nutrient-Chl *a* relationship. The inference of BNs and subsequent simulation of the BHM both showed that TP played a dominant role in determining Chl *a* concentration in both shallow and deep lakes, while TN overall played a secondary role, though depending on lake type. The effect of TN on Chl *a* tends to be more important when the lake TP is >40 μg/L. Subsequent simulations of the BHM with variable TP and TN ranges revealed variable sensitivity of the Chl *a* response to nutrients between shallow lakes, transitional lakes, and deep lakes. Deep lakes had the lowest yield Chl *a* per unit of nutrient, following transitional lakes, and shallow lakes were the highest. These results demonstrated that reduction of phosphorus is the most important way to mitigate eutrophication but a dual N and P reduction may be more effective in eutrophic lake, especially in shallow lakes. Our study showed that the lake-type nutrient-Chl *a* or lake-specific nutrient-Chl *a* relationship is to be preferred in eutrophication management. Categorization of lake according to their depth and trophic status can help lake managers specifying better targeted control objectives and effective measures to minimize the risk of eutrophication. Our established BHM based on cross-sectional data provides a flexible and reliable risk assessment framework for driving the TP and TN targets required for meeting a target Chl *a* concentration required for abating eutrophication at the level of lake type or for specific lakes. When using these models care should be taken giving that TN and TP are not fully independent predictors of phytoplankton biomass.

Funding

This work was supported by the National Natural Science Foundation of China (Nos. 41961019 and 42071131) and Yunnan Provincial Major Science and Technology Special Plan Projects (202102AE090030). EJ was supported by the TÜBITAK program BIDEB2232 (project 118C250).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2023.120182.

References

Abell, J.M., Özkundakci, D., Hamilton, D.P., 2010. Nitrogen and phosphorus limitation of phytoplankton growth in New Zealand lakes: implications for eutrophication control. Ecosystems 13 (7), 966–977. https://doi.org/10.1007/s10021-010-9367-9.

Aguilera, P., Fernández, A., Ropero, F., R, Rumí, R, Salmerón, A, 2012. Bayesian networks in environmental modelling. Environ. Model. Softw. 26, 1376–1388. https://doi.org/10.1016/j.envsoft.2011.06.004.

Canfield, D.E., Bachmann, R.W., Hoyer, M.V., Johansson, L.S., Søndergaard, M., Jeppesen, E., 2019. To measure chlorophyll or phytoplankton biovolume: an aquatic conundrum with implications for the management of lakes. Lake Reserv. Manag. 35 (2), 181–192. https://doi.org/10.1080/10402381.2019.1607958.

Cantin, A., Weyhenmeyer, G.A., Beisner, B.E., Gunn, J.M., Prairie, Y.T., Winter, J.G., 2011. Effects of thermocline deepening on lake plankton communities. Can. J. Fish. Aquat.Sci. 68 (2), 260–276. https://doi.org/10.1139/f10-138.

Carpenter, S.R., 2008. Phosphorus control is critical to mitigating eutrophication. Proc. Natl Acad. Sci. 105 (32), 11039–11040. https://doi.org/10.1073/pnas.0806112105.

Chaffin, J.D., Bridgeman, T.B., Bade, D.L., Mobilian, C.N., 2014. Summer phytoplankton nutrient limitation in Maumee Bay of Lake Erie during high-flow and low-flow years. J. Great Lakes Res. 40 (3), 524–531. https://doi.org/10.1016/j.jglr.2014.04.009.

Cha, Y., Soon Park, S., Won Lee, H., Stow, C.A., 2016. A Bayesian hierarchical approach to model seasonal algal variability along an upstream to downstream river gradient. Water Resour. Res. 52 (1), 348–357. https://doi.org/10.1002/2015wr017327.

Chen, S.H., Pollino, C.A., 2012. Good practice in Bayesian network modelling. Environ. Model. Softw. 37, 134–145. https://doi.org/10.1016/j.envsoft.2012.03.012.

Diehl, S., 2002. Phytoplankton, light, and nutrients in a gradient of mixing depths: theory. Ecology 83 (2), 386–398. https://doi.org/10.1890/0012-9658(2002)083[0386:PLANIA]2.0.CO;2.

Diehl, S., Berger, S., Ptacnik, R., Wild, A., 2002. Phytoplankton, light, and nutrients in a gradient of mixing depths: field experiments. Ecology 83 (2), 399–411. https://doi.org/10.1890/0012-9658(2002)083[0399:PLANIA]2.0.CO;2.

Dubourg, P., North, R.L., Hunter, K., Vandergucht, D.M., Abirhire, O., Silsbe, G.M., Guildford, S.J., Hudson, J.J., 2015. Light and nutrient co-limitation of phytoplankton communities in a large reservoir: Lake Diefenbaker, Saskatchewan, Canada. J. Gt. Lakes Res. 41, 129–143. https://doi.org/10.1016/j.jglr.2015.10.001.

Downing, J.A., McCauley, E., 1992. The nitrogen : phosphorus relationship in lakes. Limnol. Oceanogr. 37 (5), 936–945. https://doi.org/10.4319/lo.1992.37.5.0936.

Elser, J.J., Bracken, M.E., Cleland, E.E., Gruner, D.S., Harpole, W.S., Hillebrand, H., Ngai, J.T., Seabloom, E.W., Shurin, J.B., Smith, J.E., 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecol. Lett. 10 (12), 1135–1142. https://doi.org/10.1111/j.1461-0248.2007.01113.x.

Feki-Sahnoun, W., Hamza, A., Njah, H., Barraj, N., Mahfoudi, M., Rebai, A., Hassen, M. B., 2017. A Bayesian network approach to determine environmental factors controlling Karenia selliformis occurrences and blooms in the Gulf of Gabes, Tunisia. Harmful Algae 63, 119–132. https://doi.org/10.1016/j.hal.2017.01.013.

Filstrup, C.T., Wagner, T., Soranno, P.A., Stanley, E.H., Stow, C.A., Webster, K.E., Downing, J.A., 2014. Regional variability among nonlinear chlorophyll-phosphorus relationships in lakes. Limnol. Oceanogr. 59 (5), 1691–1703. https://doi.org/10.4319/lo.2014.59.5.1691.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. Bayesian Data Analysis, 3rd ed. Chapman and Hall/CRC. https://doi.org/10.1201/b16018.

Genkai-Kato, M., Carpenter, S.R., 2005. Eutrophication due to phosphorus recycling in relation to lake morphometry, temperature, and macrophytes. Ecology 86 (1), 210–219. https://doi.org/10.1890/03-0545.

Gonzalez Sagrario, M.A., Jeppesen, E., Goma, J., Sondergaard, M., Jensen, J.P., Lauridsen, T., Landkildehus, F., 2005. Does high nitrogen loading prevent clear-water conditions in shallow lakes at moderately high phosphorus concentrations? Freshw. Biol. 50 (1), 27–41. https://doi.org/10.1111/j.1365-2427.2004.01290.x.

Hanna, M., 1990. Evaluation of models predicting mixing depth. Can. J. Fish. Aquat.Sci. 47, 940–947. https://doi.org/10.1139/f90-108.

Ho, J.C., Michalak, A.M., Pahlevan, N., 2019. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature 574 (7780), 667–670. https://doi.org/10.1038/s41586-019-1648-7.

Huang, J., Zhang, Y., Arhonditsis, G.B., Gao, J., Chen, Q., Peng, J., 2020. The magnitude and drivers of harmful algal blooms in China's lakes and reservoirs: a national-scale characterization. Water Res. 181, 115902 https://doi.org/10.1016/j.watres.2020.115902.

Janssen, A.B.G., Janse, J.H., Beusen, A.H.W., Chang, M., Harrison, J.A., Huttunen, I., Kong, X., Rost, J., Teurlincx, S., Troost, T.A., van Wijk, D., Mooij, W.M., 2019. How to model algal blooms in any lake on earth. Curr. Opin. Environ. Sustain. 36, 1–10. https://doi.org/10.1016/j.cosust.2018.09.001.

Jeppesen, E., Canfield, D.E., Bachmann, R.W., Søndergaard, M., Havens, K.E., Johansson, L.S., Lauridsen, T.L., Sh, T., Rutter, R.P., Warren, G., Ji, G., Hoyer, M.V., 2020. Toward predicting climate change effects on lakes: a comparison of 1656 shallow lakes from Florida and Denmark reveals substantial differences in nutrient dynamics, metabolism, trophic structure, and top-down control. Inland Waters 10 (2), 197–211. https://doi.org/10.1080/20442041.2020.1711681.

Jeppesen, E., Søndergaard, M., Meerhoff, M., Lauridsen, T.L., Jensen, J.P., 2007. Shallow lake restoration by nutrient loading reduction—some recent findings and challenges ahead. Hydrobiologia 584 (1), 239–252. https://doi.org/10.1007/s10750-007-0596-7.

Jeppesen, E., Johansson, L., Sh, T., Søndergaard, M., Lauridsen, T.L., Andersen, P., 2021. Do cross-latitude and local studies give similar predictions of phytoplankton responses to warming: an analysis of monitoring data from 504 Danish lakes. Sustainability 13, 14049. https://doi.org/10.3390/su132414049.

Jessica, B.M., Hunter, J.C., 2007. Seasonal variation of phytoplankton nutrient limitation in Lake Erie. Aquat. Microb. Ecol. 48 (1), 61–71. https://doi.org/10.1016/j.ecoenv.2013.05.006.

Kolzau, S., Wiedner, C., Rucker, J., Kohler, J., Kohler, A., Dolman, A.M., 2014. Seasonal patterns of nitrogen and phosphorus limitation in four German lakes and the predictability of limitation status from ambient nutrient concentrations. PLoS One 9 (4), e96065. https://doi.org/10.1371/journal.pone.0096065.

Liang, Z., Soranno, P.A., Wagner, T., 2020. The role of phosphorus and nitrogen on chlorophyll a: evidence from hundreds of lakes. Water Res. 185, 116236 https://doi.org/10.1016/j.watres.2020.116236.

Longhi, M.L., Beisner, B.E., 2009. Environmental factors controlling the vertical distribution of phytoplankton in lakes. J. Plankton Res. 31 (10), 1195–1207. https://doi.org/10.1093/plankt/fbp065.

Maberly, S.C., Pitt, J.A., Davies, P.S., Carvalho, L., 2020. Nitrogen and phosphorus limitation and the management of small productive lakes. Inland Waters 10 (2), 159–172. https://doi.org/10.1080/20442041.2020.1714384.

Malakoff, D., 1999. Bayes offers a 'New' way to make sense of numbers. Science 286 (5444), 1460–1464. https://doi.org/10.1126/science.286.5444.1460.

Malve, O., Qian, S.S., 2006. Estimating nutrients and chlorophyll a relationships in Finnish Lakes. Environ. Sci. Technol. 40 (24), 7848–7853. https://doi.org/10.1021/es061359b.

Mellios, N.K., Moe, S.J., Laspidou, C., 2020. Using Bayesian hierarchical modelling to capture cyanobacteria dynamics in Northern European lakes. Water Res. 186, 116356 https://doi.org/10.1016/j.watres.2020.116356.

Messager, M.L., Lehner, B., Grill, G., Nedeva, I., Schmitt, O., 2016. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. Nat. Commun. 7, 13603. https://doi.org/10.1038/ncomms13603.

Morris, D.P., Lewis Jr, W.M, 1988. Phytoplankton nutrient limitation in Colorado mountain lakes. Freshw. Biol. 20 (3), 315–327. https://doi.org/10.1111/j.1365-2427.1988.tb00457.x.

Muller, B., Thoma, R., Baumann, K.B.L., Callbeck, C.M., Schubert, C.J., 2021. Nitrogen removal processes in lakes of different trophic states from on-site measurements and historic data. Aquat. Sci. 83 (2), 37. https://doi.org/10.1007/s00027-021-00795-7.

Olsen, S., Chan, F., Li, W., Zhao, S., Søndergaard, M., Jeppesen, E., 2015. Strong impact of nitrogen loading on submerged macrophytes and algae: a long-term mesocosm experiment in a shallow Chinese lake. Freshw. Biol. 60 (8), 1525–1536. https://doi.org/10.1111/fwb.12585.

Paerl, H.W., Scott, J.T., McCarthy, M.J., Newell, S.E., Gardner, W.S., Havens, K.E., Hoffman, D.K., Wilhelm, S.W., Wurtsbaugh, W.A., 2016. It takes two to tango: when and where dual nutrient (N & P) reductions are needed to protect lakes and downstream ecosystems. Environ. Sci. Technol. 50 (20), 10805–10813. https://doi.org/10.1021/acs.est.6b02575.

Pearl, J., 1985. Bayesian networks: a model of self-activated memory for evidential reasoning. In: Proceedings of the 7th Conference of the Cognitive Science Society. Irvine, CA. University of California, pp. 329–334.

Phillips, G., Pietiläinen, O.P., Carvalho, L., Solimini, A., Lyche Solheim, A, Cardoso, A.C., 2008. Chlorophyll–nutrient relationships of different lake types using a large European dataset. Aquat. Ecol. 42 (2), 213–226. https://doi.org/10.1007/s10452-008-9180-0.

Phillips, G., Willby, N, Moss, B., 2016. Submerged macrophyte decline in shallow lakes: What have we learnt in the last forty years? Aquat. Bot. 135, 37–45. https://doi.org/10.1016/j.aquabot.2016.04.004.

Poikane, S., Phillips, G., Birk, S., Free, G., Kelly, M.G., Willby, N.J., 2019. Deriving nutrient criteria to support 'good' ecological status in European lakes: An empirically based approach to linking ecology and management. Sci. Total. Environ. 650, 2074–2084. https://doi.org/10.1016/j.scitotenv.2018.09.350. Pt 2.

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. https://doi.org/10.1007/s10021-010-9380-z.

Qin, B., Paerl, H.W., Brookes, J.D., Liu, J., Jeppesen, E., Zhu, G., Zhang, Y., Xu, H., Shi, K., Deng, J., 2019. Why lake Taihu continues to be plagued with cyanobacterial blooms through 10 years (2007–2017) efforts. Sci. Bull. 64 (6), 354–356. https://doi.org/10.1016/j.scib.2019.02.008.

Qin, B., Zhou, J., Elser, J.J., Gardner, W.S., Deng, J., Brookes, J.D., 2020. Water depth underpins the relative roles and fates of nitrogen and phosphorus in lakes. Environ. Sci. Technol. 54 (6), 3191–3198. https://doi.org/10.1021/acs.est.9b05858.

Salvatier, J., Wiecki, T.V., Fonnesbeck, C., 2016. Probabilistic programming in Python using PyMC3. PeerJ Comput. Sci. 2 https://doi.org/10.7717/peerj-cs.55.

Saunders, D.L., Kalff, J., 2001. Denitrification rates in the sediments of Lake Memphremagog, Canada–USA. Water Res. 35 (8), 1897–1904. https://doi.org/10.1016/S0043-1354(00)00479-6.

Schindler, D.W., 2012. The dilemma of controlling cultural eutrophication of lakes. Proc. Biol. Sci. 279 (1746), 4322–4333. https://doi.org/10.1098/rspb.2012.1032.

Schindler, D.W., Carpenter, S.R., Chapra, S.C., Hecky, R.E., Orihel, D.M., 2016. Reducing Phosphorus to Curb Lake Eutrophication is a Success. Environ. Sci. Technol. 50 (17), 8923–8929. https://doi.org/10.1021/acs.est.6b02204.

Scutari, M., 2010. Learning Bayesian Networks with the bnlearn R Package. J. Stat. Softw. 35 (3), 1–22. https://doi.org/10.18637/jss.v035.i03.

Søndergaard, M., Jensen, J.P., Jeppesen, E., 2003. Role of sediment and internal loading of phosphorus in shallow lakes. Hydrobiologia 506 (1), 135–145. https://doi.org/10.1023/B:HYDR.0000008611.12704.dd.

Søndergaard, M., Jeppesen, E., Lauridsen, T.L., Skov, C., Van Nes, E.H., Roijackers, R., Lammens, E., Portielje, R.O.B., 2007. Lake restoration: successes, failures and long-term effects. J. Appl. Ecol. 44 (6), 1095–1105. https://doi.org/10.1111/j.1365-2664.2007.01363.x.

Søndergaard, M., Lauridsen, T.L., Johansson, L.S., Jeppesen, E., 2017. Nitrogen or phosphorus limitation in lakes and its impact on phytoplankton biomass and submerged macrophyte cover. Hydrobiologia 795 (1), 35–48. https://doi.org/10.1007/s10750-017-3110-x.

Soranno, P.A., Bacon, L.C., Beauchene, M., Bednar, K.E., Bissell, E.G., Boudreau, C.K., Boyer, M.G., Bremigan, M.T., Carpenter, S.R., Carr, J.W., Cheruvelil, K.S., Christel, S. T., Claucherty, M., Collins, S.M., Conroy, J.D., Downing, J.A., Dukett, J., Fergus, C. E., Filstrup, C.T., Funk, C., Gonzalez, M.J., Green, L.T., Gries, C., Halfman, J.D., Hamilton, S.K., Hanson, P.C., Henry, E.N., Herron, E.M., Hockings, C., Jackson, J.R., Jacobson-Hedin, K., Janus, L.L., Jones, W.W., Jones, J.R., Keson, C.M., King, K.B.S., Kishbaugh, S.A., Lapierre, J.F., Lathrop, B., Latimore, J.A., Lee, Y., Lottig, N.R., Lynch, J.A., Matthews, L.J., McDowell, W.H., Moore, K.E.B., Neff, B.P., Nelson, S.J., Oliver, S.K., Pace, M.L., Pierson, D.C., Poisson, A.C., Pollard, A.I., Post, D.M., Reyes, P.O., Rosenberry, D.O., Roy, K.M., Rudstam, L.G., Sarnelle, O., Schuldt, N.J., Scott, C.E., Skaff, N.K., Smith, N.J., Spinelli, N.R., Stachelek, J.J., Stanley, E.H., Stoddard, J.L., Stopyak, S.B., Stow, C.A., Tallant, J.M., Tan, P.N., Thorpe, A.P., Vanni, M.J., Wagner, T., Watkins, G., Weathers, K.C., Webster, K.E., White, J.D., Wilmes, M.K., Yuan, S., 2017. LAGOS-NE: a multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of US lakes. GigaScience 6 (12), 1–22. https://doi.org/10.1093/gigascience/gix101.

USEPA, 2009. National Lakes Assessment: A Collaborative Survey of the Nation's Lakes. U.S. Environmental Protection Agency, Washington, D.C..

Uusitalo, L., 2007. Advantages and challenges of Bayesian networks in environmental modelling. Ecol. Model. 203 (3-4), 312–318. https://doi.org/10.1016/j.ecolmodel.2006.11.033.

Wu, S., Wu, Z., Liang, Z., Liu, Y., Wang, Y., 2019. Denitrification and the controlling factors in Yunnan Plateau Lakes (China): exploring the role of enhanced internal nitrogen cycling by algal blooms. J. Environ. Sci. 76, 349–358. https://doi.org/10.1016/j.jes.2018.05.028 (China).

Yang, R., Fan, X., Zhao, L., Yang, K., 2023. Identification of major environmental factors driving phytoplankton community succession before and after the regime shift of Erhai Lake, China. Ecol. Indic. 146 https://doi.org/10.1016/j.ecolind.2023.109875.

Zhang, Y., Jeppesen, E., Liu, X., Qin, B., Shi, K., Zhou, Y., Thomaz, S.M., Deng, J., 2017. Global loss of aquatic vegetation in lakes. Earth Sci. Rev. 173, 259–265. https://doi.org/10.1016/j.earscirev.2017.08.013.

Zhao, L., Cheng, S., Sun, Y., Zou, R., Ma, W., Zhou, Q., Liu, Y., 2021. Thermal mixing of Lake Erhai (Southwest China) induced by bottom heat transfer: Evidence based on observations and CE-QUAL-W2 model simulations. J. Hydrol. 603. https://doi.org/10.1016/j.jhydrol.2021.126973.

Zhu, L., Shi, W., Zhou, J., Yu, J., Kong, L., Qin, B., 2021. Strong turbulence accelerates sediment nitrification-denitrification for nitrogen loss in shallow lakes. Sci. Total. Environ. 761, 143210 https://doi.org/10.1016/j.scitotenv.2020.143210.