

# Chlorophyll–nutrient relationships of different lake types using a large European dataset

G. Phillips · O.-P. Pietiläinen · L. Carvalho ·  
A. Solimini · A. Lyche Solheim · A. C. Cardoso

Published online: 25 April 2008  
© Springer Science+Business Media B.V. 2008

**Abstract** In Europe there is a renewed focus on relationships between chemical determinands and ecological impact as a result of the Water Framework Directive (WFD). In this paper we use regression analysis to examine the relationship of growing season mean chlorophyll *a* concentration with total phosphorus and total nitrogen using summary data from over 1,000 European lakes. For analysis, lakes were grouped into types with three categories of mean depth, alkalinity and humic content. The lakes were also divided into broad geographic regions covering Atlantic, Northern, Central/Baltic and for some types the Mediterranean areas of Europe. Chlorophyll *a* was found to be significantly related to both total phosphorus and total nitrogen, although total phosphorus was almost always found to be the

best predictor of chlorophyll. Different nutrient chlorophyll relationships were found for lakes according to mean depth and alkalinity, although no significant effect of geographic region or humic content was found for the majority of lake types. We identified three groups of lakes with significantly different responses. Deep lakes had the lowest yield of chlorophyll per unit of nutrient, low and moderate alkalinity shallow lakes the highest and high alkalinity lakes were intermediate. We recommend that the regression models provided for these three lake groups should be used for lake management in Europe and discuss the limitations of such models.

**Keywords** Eutrophication · Lakes · Chlorophyll · Nitrogen · Phosphorus · Regression model · Water Framework Directive · Lake types

---

G. Phillips (✉)  
Environment Agency (EA), Reading, UK  
e-mail: Geoff.phillips@environment-agency.gov.uk

O.-P. Pietiläinen  
Finnish Environment Institute (SYKE), Helsinki, Finland

L. Carvalho  
Centre for Ecology and Hydrology (CEH), Edinburgh,  
UK

A. Solimini · A. Lyche Solheim · A. C. Cardoso  
Joint Research Centre (JRC), Ispra, Italy

A. Lyche Solheim  
Norwegian Institute for Water Research (NIVA), Oslo,  
Norway

## Introduction

The introduction of the Water Framework Directive (2000/60/EC) in Europe has focused the attention of many lake managers and scientists on relationships between ecological status and anthropogenic pressures. For lakes, nutrients remain one of the key pressures influencing European lakes and one of the most obvious and well known effects that need to be quantified is the impact that elevated nutrients have on phytoplankton biomass.

Chlorophyll *a* concentration is a widely used measure of phytoplankton biomass and since the founding papers of Sakamoto (1966) and Dillon and Rigler (1974) there have been many studies showing strong empirical links between chlorophyll and nutrients, particularly phosphorus (cf. Prairie et al. 1989). The most widely known relationships are those from the OECD study (Vollenweider and Kerekes 1980) and these are often used to underpin decisions about lake management (Ryding and Rast 1989). Reynolds (1980) has highlighted the limitations of such models for individual lake predictions, but their value to lake managers, who need to make decisions with limited data, is undeniable.

One problem with the published relationships is the substantial variability of the parameters generated from different regression analyses. These differences are due to the many influences on the nutrient chlorophyll relationship. Regressions drawn from specific lake districts are particularly appropriate for making predictions about lakes within those districts, but they may be of less use for lakes in other areas. Conversely, large datasets will be influenced by many and various factors, all of which will appear as increased variation. In this paper we present the results of the analysis of a large dataset drawn from a wide variety of national European data archives. In particular we explore the variation of nutrient chlorophyll relationships in different lake types with a view to providing a set of regression equations that can be used to underpin lake nutrient management decisions for European lakes.

## Materials and methods

Data for this study were gathered from national data archives from individual European countries, as a basis for studying relationships between biological and chemical elements in European lakes as part of the EU project REBECCA (Moe et al. 2008). As part of the WFD Common Implementation Strategy (Van de Bund et al. 2004) lakes have been divided into six geographic regions. Individual sample data were available from 16 countries, spanning a period from 1988 to 2004. Five of the six regions were represented, although relatively few data were available from the Mediterranean region, and the dataset was dominated by data from countries in northern, and to a lesser extent central Europe.

Chlorophyll *a* (Chl), total phosphorus (TP) and total nitrogen (TN) data were summarised for each lake as a single mean value for a common “growing season” of April–September. In a large dataset drawn from many different sources, different sampling frequencies could influence summary statistics. To minimise noise, only lakes with more than three sampling occasions during this growing season were selected, and data were summarised by site year and month and the above three determinands matched prior to calculation of a single mean for each lake. This ensured that only matched summary data for Chl, TP and TN contributed to the final mean value. Data for all available years and sites within a lake contributed to a final single lake average value for each of Chl, TP and TN, thereby obtaining only one single value per lake for each parameter in the regression analyses.

This resulted in a dataset of 1,138 lakes, substantially larger than any other reported in the literature that we are aware of (Table 2). As pointed out by Prairie et al. (1989) this brings with it additional uncertainties, caused by different geographical regions, climate and sampling strategies. We will examine the effect of region and lake types, and thus assigned each lake into one of nine core types based on alkalinity and mean depth and three sub-types based on colour (Table 1). Data for alkalinity, depth and colour were taken from national data archives, collated by Moe et al. (2008). These types were identified by a range of European experts during the WFD intercalibration

**Table 1** Details of lake types used to categorise lakes

Type	Abbreviation	Determinand
Alkalinity types		
High	HA	Alkalinity > 1.0 mEq l <sup>-1</sup>
Moderate	MA	Alkalinity 0.2–1.0 mEq l <sup>-1</sup>
Low	LA	Alkalinity < 0.2 mEq l <sup>-1</sup>
Depth types		
Deep	D	Mean depth > 15.0 m
Shallow	S	Mean depth 3.0–15.0 m
Very shallow	VS	Mean depth < 3.0 m
Colour types		
Polyhumic	VH	Colour > 90 mg Pt l <sup>-1</sup>
Humic	H	Colour 30–90 mg Pt l <sup>-1</sup>
Clear	L	Colour < 30 mg Pt l <sup>-1</sup>

process and represent the best available agreed typology that is likely to minimise natural biological variation in European lakes.

Chlorophyll nutrient relationships, denoted as  $\text{Chl} = f(\text{TP})$ ,  $\text{Chl} = f(\text{TN})$  and  $\text{Chl} = f(\text{TP}, \text{TN})$ , were derived using linear least squares regression. Prior to analysis all data were logarithmically transformed (base 10) to ensure their homogeneity of variance, as is common practice for data of this type. We investigated the effects of categorical variables lake type, geographic region and humic content on these relationships using univariate general linear modelling (GLM). The model was run with Chl as dependent variable, TP as covariate and each of the above categorical variables as fixed factors. Analysis of variance (ANOVA) was used to test the significance of the factor on the slope, intercept and their interactions. Type III sums of squares were used to test significance, which had the advantage that they are not dependent on the frequency of observations in each group (Field 2005).

Having identified which type factors significantly influenced the regression, we compared the resulting regression equations to identify which types had similar regression coefficients. Lake types with similar coefficients were then re-grouped and the general linear model (GLM) repeated to check that the type factor was no longer significantly influencing the regression relationship. Thus, we were able to identify the key factors influencing the nutrient Chl relationship and identify appropriate groups of lake types.

Prior to analysis, we investigated the general shape of the relationships between variables using LOWESS techniques (Cleveland 1979). This enabled us to restrict regression to a range of nutrient concentrations where the data were linearly related. For  $\text{Chl} = f(\text{TP})$  we therefore restricted regression to data where  $\text{TP} < 100 \mu\text{g l}^{-1}$  and for  $\text{Chl} = f(\text{TN})$  where  $\text{TN} < 1,700 \mu\text{g l}^{-1}$ . For  $\text{Chl} = f(\text{TP}, \text{TN})$  no restrictions were made as our purpose was to identify if the use of both nutrients provided an effective model for the whole data range. We tested the significance of including a second predictor variable using ANOVA. As TP and TN might be correlated, we tested for the variance inflation factor (VIF), which is a good indicator of whether one predictor has a strong linear relationship with other predictors. Where the VIF was  $< 10$

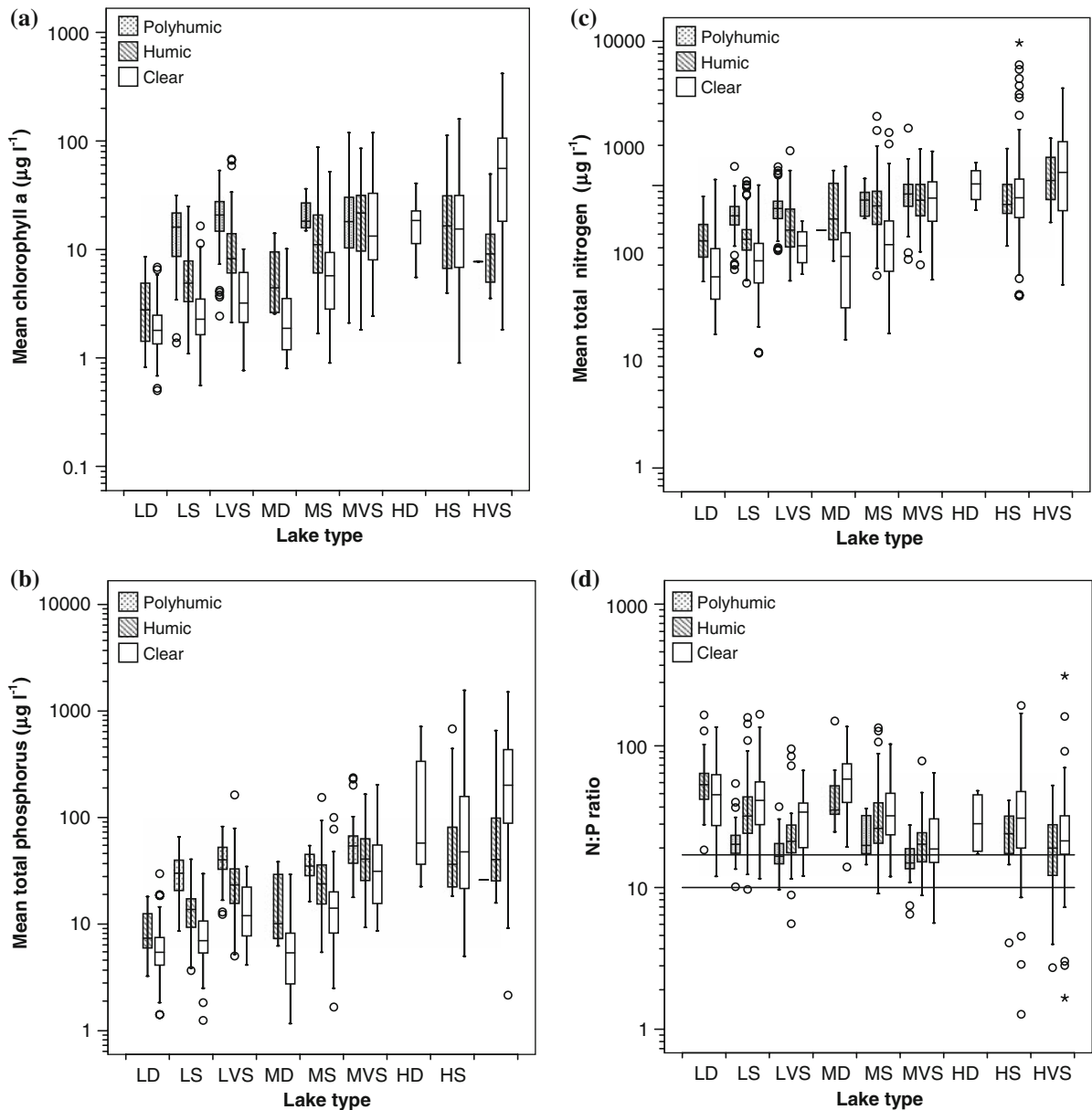
we concluded that predictor variables were not strongly related (Myers 1990).

We investigated the relative power of TP and TN in predicting Chl for different lake types using stepwise multiple regression, comparing regression coefficients as values standardised by standard deviation units ( $\beta$  coefficients) to make them independent of scale. Prior to analysis, the data were screened by removing samples with reported values of Chl or TP of  $< 1.0 \mu\text{g l}^{-1}$  as we considered these were likely to be errors in the source data. We also excluded results where the Chl:TP ratio was  $> 1.0$ , which was the 99th percentile of the whole dataset.

To explore the effect of TN:TP ratios, proposed by several studies as important factors influencing the relationship between nutrients and chlorophyll (Nürnberg and Shaw 1999), we grouped data into five TN:TP categories ( $\leq 10$ ,  $>10 \leq 17$ ,  $>17 \leq 25$ ,  $>25 \leq 50$ ,  $>50$ ). We chose the first two categories as they are widely reported values at which nitrogen limitation of phytoplankton biomass may occur (Forsberg and Ryding 1980); the others are convenient categorical divisions. Analysis was carried out using SPSS v14.

## Results

After data screening  $> 1,000$  typed pairs of data were available for analysis (1129 TP vs. Chl; 1138 TN vs. Chl and 1077 TN and TP vs. Chl). Chlorophyll, phosphorus and nitrogen concentrations ranged over two orders of magnitude, increasing markedly as lakes get shallower and to a lesser extent as alkalinity increases (Fig. 1a–c). In low and moderate alkalinity lakes there was also a marked effect of humic content with the highest concentrations of all these parameters in polyhumic lakes and the lowest in clear water lakes. In high alkalinity lakes, however, the effect of humic content was either non-existent or even reversed in high alkalinity very shallow lakes. The ratio of TN:TP ranged from  $< 2$  to  $> 100$ , with different ranges characteristic of each lake type (Fig. 1d). TN:TP was highest in low and moderate alkalinity, deep lakes and became progressively lower as depth decreased and alkalinity increased. Polyhumic and humic lakes nearly always had lower TN:TP ratios than clear water lakes, and a substantial proportion of lakes in the coloured, very shallow lake



**Fig. 1** Box plots showing type-specific ranges of seasonal (April–September) mean values for (a) chlorophyll *a* ( $\mu\text{g l}^{-1}$ ), (b) total phosphorus ( $\mu\text{g l}^{-1}$ ), (c) total nitrogen ( $\mu\text{g l}^{-1}$ ), (d) total nitrogen to total phosphorus ratio in lakes of different humic class. Lakes split into three alkalinity classes (low L, moderate M and high H) and three depth classes (deep D,

shallow S, very shallow VS). Boxes represent polyhumic (dotted), humic (striped) and clear water (clear) lakes and show median, interquartile range and outliers. Horizontal lines (d) represent ratios of 10 and 17 marking probable region of N limitation and NP co-limitation respectively

categories had values below 17 and a minority of lakes below 10. The effect of colour on the TN:TP ratio was confirmed by two way ANOVA, which showed significant effects of both lake type

( $F = 12.1$   $P < 0.001$ ) and humic type ( $F = 9.7$   $P < 0.001$ ) with post hoc Bonferroni tests showing significant differences in mean N:P ratios for all comparisons of humic categories.

## Regression relationships

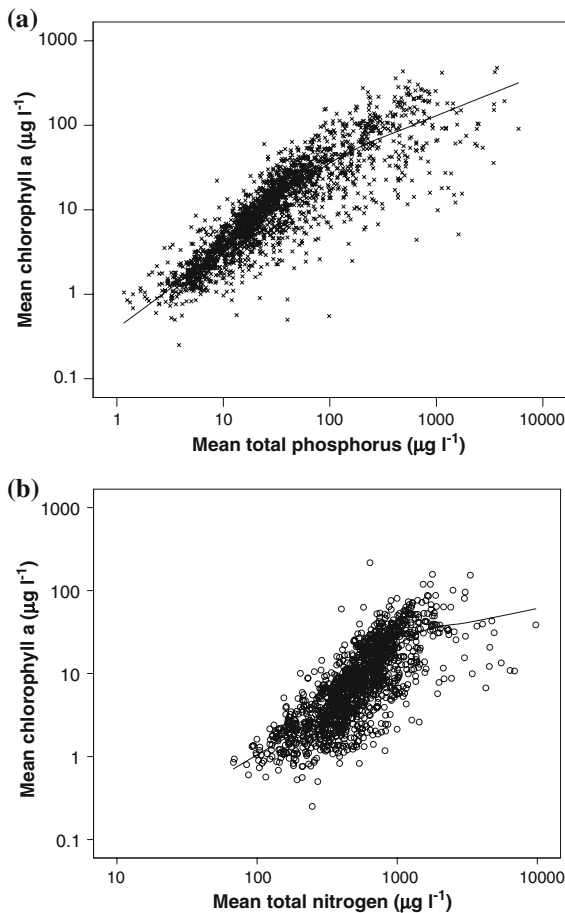
Scatter plots for the relationships  $\text{Chl} = f(\text{TP})$  and  $\text{Chl} = f(\text{TN})$  are shown in Fig. 2. The LOESS fitted line shows non-linearity in both relationships. There is slight indication of a sigmoidal response, but the mostly marked non-linearity is the reduction in slope at high nutrient levels. This is most marked for TN, but also occurs for TP. The upper point of inflection for  $\text{Chl} = f(\text{TP})$  occurs at ca.  $\text{TP} = 100 \mu\text{g l}^{-1}$  and for  $\text{Chl} = f(\text{TN})$  at  $\text{TN} = 1,700 \mu\text{g l}^{-1}$ . Asymptotic behaviour of  $\text{Chl} = f(\text{TP})$  is widely reported, occurring typically at around this TP concentration (Prairie et al. 1989; Canfield 1983), although in our case the LOESS fits do not give as clear an indication of a plateau in chlorophyll as do those of others. However, because

of this we restricted the range of TP and TN concentrations in our least square regressions to concentrations of  $< 100 \mu\text{g l}^{-1}$  (TP) and  $< 1,700 \mu\text{g l}^{-1}$  (TN) (Fig. 3).

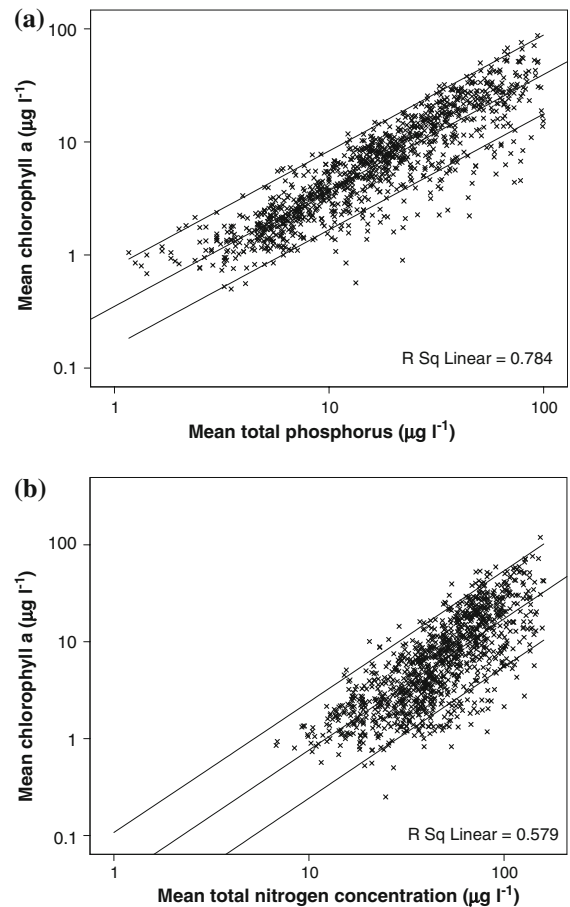
The regressions of all lakes taken together regardless of type were the following:

1.  $\text{Log}_{10} \text{Chl} = -0.455(\pm 0.020) + 1.026(\pm 0.016) \text{Log}_{10} \text{TP}$   $R^2 = 0.78$  (TP  $< 100 \mu\text{g l}^{-1}$ )
2.  $\text{Log}_{10} \text{Chl} = -2.828(\pm 0.093) + 1.355(\pm 0.035) \text{Log}_{10} \text{TN}$   $R^2 = 0.58$  (TN  $< 1,700 \mu\text{g l}^{-1}$ )
3.  $\text{Log}_{10} \text{Chl} = -1.028(\pm 0.083) + 0.792(\pm 0.027) \text{Log}_{10} \text{TP} + 0.324(\pm 0.039) \text{Log}_{10} \text{TN}$   $R^2 = 0.77$

All regressions were significant ( $P < 0.001$ ) and the coefficients of determination ( $R^2$ ) were similar to those reported in other published studies (Table 2). Scatter plots with the regressions are shown in Fig. 3.



**Fig. 2** Relationship between growing season chlorophyll *a* concentration ( $\mu\text{g l}^{-1}$ ) and (a) total phosphorus concentration ( $\mu\text{g l}^{-1}$ ), (b) total nitrogen concentration ( $\mu\text{g l}^{-1}$ ). Solid line is the LOESS fitted curve



**Fig. 3** Relationship between growing season chlorophyll *a* concentration ( $\mu\text{g l}^{-1}$ ) and (a) total phosphorus concentration ( $\mu\text{g l}^{-1}$ ) (b) total nitrogen concentration ( $\mu\text{g l}^{-1}$ ). Lines are best fit regression and 95% confidence limits

**Table 2** Regression equations from this study (in *italics*) compared to those taken from the literature for  $\log_{10}$  chlorophyll *a* concentration as a function of total phosphorus (TP) and total nitrogen (TN)

Reference	Constant	Slope		$R^2$	Comment	Predicted Chlorophyll
		Log TP	Log TN			
Chlorophyll $a$ f(TP)						
OECD (1982)	−0.432	0.79		0.77		5.4
Classen (1980) (OECD shallow lakes and reservoirs)	−0.268	0.720		0.76		6.3
White (1983)	−0.638	0.940		0.45		6.5
<i>This study (deep lakes)</i>	−0.286	0.776		0.65		8.2
Rast and Lee (1978)	−0.260	0.760		0.59		8.2
Prairie et al. (1989)	−0.390	0.874		0.69		9.1
Vollenweider (1976)	−0.432	0.910		0.76		9.4
Havens and Nürnberg (2004)	−0.156	0.738		0.60	Humic	9.6
Nürnberg (1996)	−0.250	0.799		0.64		9.6
Havens and Nürnberg (2004)	−0.240	0.813		0.59	Clear	10.4
<i>This study (High alkalinity shallow and very shallow lakes)</i>	−0.306	0.868		0.52		10.8
Dillon and Rigler (1974)	−1.136	1.449		0.92		12.6
<i>This study (all lakes)</i>	−0.455	1.026		0.78		13.5
Jones and Bachmann (1976)	−1.090	1.460				14.8
<i>This study (Low and moderate alkalinity shallow and very shallow lakes)</i>	−0.528	1.108		0.81		15.2
Seip et al. (2000)	−0.443	1.123		0.93		19.5
Chlorophyll $a$ f(TN)						
Nürnberg (1996)	−2.180		1.114	0.38		12.5
Prairie et al. (1989)	−3.131		1.445	0.69		13.2
Prairie et al. (1989)	−2.888		1.371		N:P 20–30	14.0
<i>This study (all lakes)</i>	−2.828		1.355	0.58		14.4
<i>This study (lakes N:P 10–17)</i>	−1.761		1.034	0.53	N:P < 10	19.1
White (1983)	−2.699		1.410	0.74	N:P < 17	28.1
Prairie et al. (1989)	−1.627		1.072		N:P < 10	33.6
Chlorophyll $a$ f(TP, TN)						
Smith (1982)	−1.517	0.653	0.548	0.76		12.7
Prairie et al. (1989)	−2.213	0.517	0.838	0.81		11.2
<i>This study (all lakes)</i>	−1.135	0.896	0.322	0.80		15.7

To facilitate comparison of non-linear functions the predicted chlorophyll concentration ( $\mu\text{g l}^{-1}$ ) is given when TP = 35  $\mu\text{g l}^{-1}$  and TN = 875  $\mu\text{g l}^{-1}$  (N:P ratio of 25)

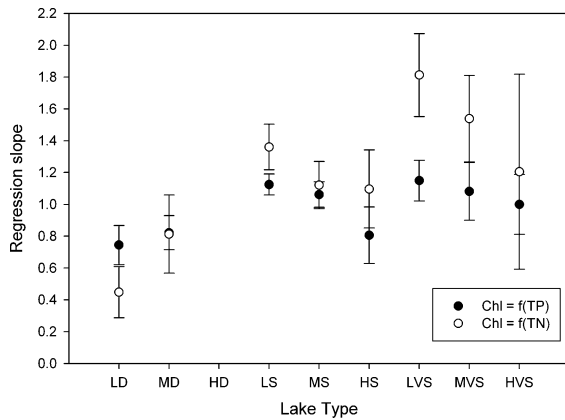
Within the linear range of these functions, TP was a better predictor than TN, although over the full data range the combined Chl-f(TN, TP) model gave almost as good a prediction.

We also tested the use of a Chl = f(TN, TP) model restricted to the linear portion of both the TP and TN relationships.

$$4. \log_{10} \text{Chl} = -1.135(\pm 0.026) + 0.896(\pm 0.026) \log_{10} \text{TP} + 0.322(\pm 0.039) \log_{10} \text{TN} \quad R^2 = 0.80$$

(TP < 100  $\mu\text{g l}^{-1}$ , TN < 1700  $\mu\text{g l}^{-1}$ )

In both models collinearity was found to be low (VIF 2.6 model 3 and 2.5 model 4). For both models 3 and 4 the addition of TN provided a small, but significant ( $F = 68.2$ ,  $P < 0.001$  model 3 and 67.4  $P < 0.001$  model 4) increase in predictive power. The standardised  $\beta$  regression coefficients for TP were, however, much higher than those for TN ( $\beta_{\text{TP}} = 0.71$   $\beta_{\text{TN}} = 0.20$  model 3;  $\beta_{\text{TP}} = 0.75$   $\beta_{\text{TN}} = 0.18$  model 4) indicating that TP was the more influential variable.



**Fig. 4** Change in regression slope for  $\text{Chl} = f(\text{TP})$  and  $\text{Chl} = f(\text{TN})$  relationships in lakes of different types

There is a significant degree of scatter in these data and we now explore whether dividing lakes into types can account for this variation and thus provide more useful regression models. The general linear model (GLM) demonstrated that while geographic region did not influence the regressions, lake type was a significant factor in the regression models; so we repeated the regression analysis on lakes split by type (Table 3). Few of the resulting  $R^2$  were higher than those obtained for the whole

dataset, but significantly different regression parameters were found. We give particular attention to the differences in slope (Fig. 4). For a log–log relationship this characterises the relative rate of increase of two variables and thus the highest slopes indicate lake types where there is the greatest positive response to TP or TN. Our all lake relationships had slopes for TP close to 1, slightly higher for TN, but in several cases the type-specific slopes were significantly different. Differences were greatest for  $\text{Chl} = f(\text{TN})$ , but for both  $\text{Chl} = f(\text{TP})$  and  $\text{Chl} = f(\text{TN})$ , deep lakes had significantly lower slopes. In general the chlorophyll response to nutrients increases significantly as depth decreases and for shallow and very shallow lakes it decreases with alkalinity. These results suggest that predictions made without reference to type may significantly over- or under-estimate chlorophyll concentrations. Type-specific relationships seem to offer the best approach for TN, but for TP we have divided lakes into three groups: (1) deep lakes, (2) moderate and low alkalinity shallow and very shallow lakes, and (3) high alkalinity shallow and very shallow lakes. Within these groups the GLM confirmed that lake type did not significantly influence the regression and the following equations

**Table 3** Regression equations for relationship between mean growing season chlorophyll  $a$  ( $\mu\text{g l}^{-1}$ ) and total phosphorus ( $\mu\text{g l}^{-1}$ ) and total nitrogen ( $\mu\text{g l}^{-1}$ ) for different European lake types (\*\*\*)  $P < 0.001$ , \*  $P < 0.05$

Type	Equation	$R^2$	$n$
High alk	Not significant	0.334	10
Deep	Too few data	0.75*	7
High alk	$\text{Log}_{10} \text{Chl} = -0.216(\pm 0.13) + 0.806(\pm 0.09) \text{Log}_{10} \text{TP}$	0.42***	115
Shallow	$\text{Log}_{10} \text{Chl} = -2.177(\pm 0.35) + 1.096(\pm 0.12) \text{Log}_{10} \text{TN}$	0.51***	78
High alk	$\text{Log}_{10} \text{Chl} = -0.521(\pm 0.15) + 1.000(\pm 0.09) \text{Log}_{10} \text{TP}$	0.71***	48
Very shallow	$\text{Log}_{10} \text{Chl} = -2.575(\pm 0.87) + 1.205(\pm 0.30) \text{Log}_{10} \text{TN}$	0.37***	29
Mod alk	$\text{Log}_{10} \text{Chl} = -0.243(\pm 0.05) + 0.822(\pm 0.05) \text{Log}_{10} \text{TP}$	0.86***	41
Deep	$\text{Log}_{10} \text{Chl} = -1.609(\pm 0.31) + 0.813(\pm 0.12) \text{Log}_{10} \text{TN}$	0.53***	42
Mod alk	$\text{Log}_{10} \text{Chl} = -0.434(\pm 0.05) + 1.062(\pm 0.04) \text{Log}_{10} \text{TP}$	0.77***	201
Shallow	$\text{Log}_{10} \text{Chl} = -2.158(\pm 0.19) + 1.091(\pm 0.07) \text{Log}_{10} \text{TN}$	0.54***	194
Mod alk	$\text{Log}_{10} \text{Chl} = -0.501(\pm 0.15) + 1.081(\pm 0.09) \text{Log}_{10} \text{TP}$	0.57***	109
Very shallow	$\text{Log}_{10} \text{Chl} = -3.189(\pm 0.40) + 1.538(\pm 0.14) \text{Log}_{10} \text{TN}$	0.53***	115
Low alk	$\text{Log}_{10} \text{Chl} = -0.283(\pm 0.05) + 0.745(\pm 0.06) \text{Log}_{10} \text{TP}$	0.50***	146
Deep	$\text{Log}_{10} \text{Chl} = -0.783(\pm 0.20) + 0.448(\pm 0.08) \text{Log}_{10} \text{TN}$	0.18***	146
Low alk	$\text{Log}_{10} \text{Chl} = -0.561(\pm 0.04) + 1.125(\pm 0.03) \text{Log}_{10} \text{TP}$	0.77***	344
Shallow	$\text{Log}_{10} \text{Chl} = -2.866(\pm 0.19) + 1.361(\pm 0.07) \text{Log}_{10} \text{TN}$	0.49***	363
Low alk	$\text{Log}_{10} \text{Chl} = -0.596(\pm 0.10) + 1.149(\pm 0.06) \text{Log}_{10} \text{TP}$	0.74***	114
Very shallow	$\text{Log}_{10} \text{Chl} = -3.904(\pm 0.36) + 1.812(\pm 0.13) \text{Log}_{10} \text{TN}$	0.61***	123



(all significant  $P < 0.001$ ) represent what we consider are appropriate models for these lake groups.

5. Low and moderate alkalinity ( $\leq 1$  mEq  $\text{l}^{-1}$ ), shallow (3–15 m mean depth) and very shallow ( $< 3$  m mean depth) lakes:

$$\text{Log}_{10} \text{Chl} = -0.528(\pm 0.03) + 1.108(\pm 0.02) \text{Log}_{10} \text{TP} \quad R^2 = 0.81$$

6. High alkalinity ( $> 1$  mEq  $\text{l}^{-1}$ ), shallow (3–15 m mean depth) and very shallow ( $< 3$  m mean depth) lakes:

$$\text{Log}_{10} \text{Chl} = -0.306(\pm 0.10) + 0.868(\pm 0.07) \text{Log}_{10} \text{TP} \quad R^2 = 0.52$$

7. All deep lakes ( $> 15$  m mean depth):

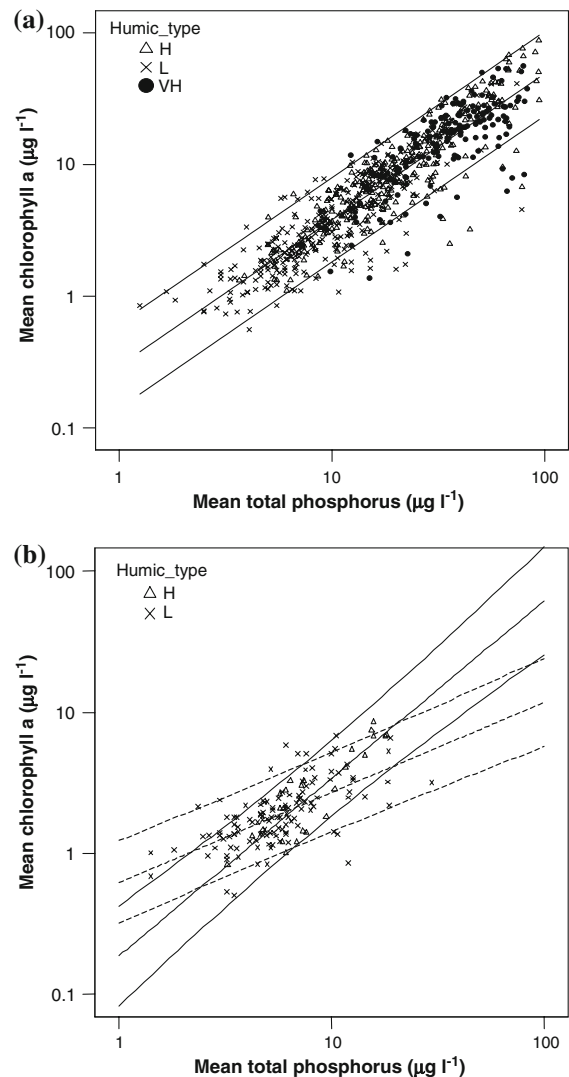
$$\text{Log}_{10} \text{Chl} = -0.286(\pm 0.04) + 0.776(\pm 0.041) \text{Log}_{10} \text{TP} \quad R^2 = 0.65$$

Colour from humic compounds is another factor which has been reported to influence the Chlorophyll nutrient relationship (Havens and Nürnberg 2004). In our data we found that TP, TN and Chl were higher in humic lakes, but GLM showed no significant effect of colour on  $\text{Chl} = f(\text{TN})$  in any lake types. For  $\text{Chl} = f(\text{TP})$  only one lake type (low alkalinity deep lakes) showed a humic type effect with a significantly higher slope and intercept in the humic lakes than in the clear water lakes (Fig. 5). The regression equations were both significant ( $P < 0.001$ ) and are given below:

8. LAD Humic:  $\text{Log}_{10} \text{Chl} = -0.733(\pm 0.14) + 1.261(\pm 0.15) \text{Log}_{10} \text{TP} \quad R^2 = 0.75$   
 9. LAD Clear:  $\text{Log}_{10} \text{Chl} = -0.208(\pm 0.05) + 0.639(\pm 0.07) \text{Log}_{10} \text{TP} \quad R^2 = 0.42$

#### Effect of TN:TP ratio on regression

Lakes were also grouped in categories based on TN:TP ratio to investigate the effects of this on the regression relationships. This ratio should indicate the potential for nitrogen or phosphorus limitation, at least at the seasonal scale, and we thus might expect that where TN:TP is low, nitrogen may limit algal growth and thus TN would provide a better prediction of chlorophyll. Conversely when TN:TP is high, phosphorus is more likely to limit growth and TP would be the better predictor. To an extent we found

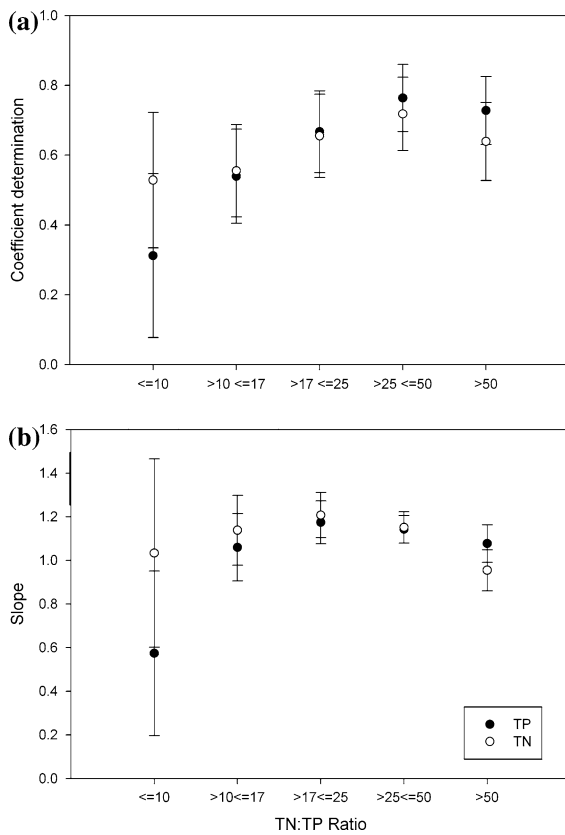


**Fig. 5** Scatter plot and regression lines ( $\pm 95\%$  confidence intervals) for  $\text{Chl} = f(\text{TP})$  for lakes of different humic content in (a) low and moderate alkalinity shallow and very shallow lakes (solid line all humic types), (b) low alkalinity deep lakes (solid line clear water (L), dotted line humic (H) types)

these predictions were true. Where TN:TP is  $< 10$ , TN provided a better prediction of Chl with a higher  $R^2$  and regression slope (Fig. 6). As TN:TP increased this difference is rapidly reduced, with similar  $R^2$  and regressions slopes for either TN or TP. The highest category (TN:TP  $> 50$ )  $\text{Chl} = f(\text{TP})$  provides a better fit and has a higher slope than  $\text{Chl} = f(\text{TN})$ , indicating the clear dominance of phosphorus limitation.

Finally, using stepwise multiple regression, we investigate the relative predictive power of TP and TN if lakes were split into alkalinity, depth and





**Fig. 6** Changes in (a) the coefficient of determination ( $R^2$ ) and (b) the slope of  $\text{Chl} = f(\text{TP})$  and  $\text{Chl} = f(\text{TN})$  for lakes grouped by total nitrogen to total phosphorus ratios (TN:TP)

humic categories. Table 4 shows the resulting values for  $R^2$  together with the standardised  $\beta$  regression coefficients (in standard deviation units), which can be used to compare the relative importance of each

predictor variable independent of scale. Stepwise regression selected  $\text{Chl} = f(\text{TP})$  for the majority of lake types, but for those lake types with the lowest TN:TP ratios (humic HAVS and polyhumic MAVS)  $\text{Chl} = f(\text{TN})$  was the preferred model. In 7 of the 19 types,  $\text{Chl} = f(\text{TP}, \text{TN})$  was selected and for three of these (polyhumic LAS, polyhumic LAVS and humic LAVS) the  $\beta$  coefficients were very similar indicating that both TP and TN had equal predictive power. All these lake types had relatively low TN:TP ratios where it might be expected that nitrogen limitation would occur. The remaining types either had TP as the single predictor variable, or if both TP and TN were selected, the  $\beta$  coefficients suggested that TN played a minor role in the model.

## Discussion

There is a relatively wide range of parameter values available in the literature relating to regression models which can be used to estimate chlorophyll *a* concentration from phosphorus (Table 2). Several studies (Prairie et al. 1989; White 1983) report values that generally predict lower chlorophyll values than those we found when we grouped all lakes together. These include the widely used equations from the OECD study (Vollenweider and Kerekes 1980). Other studies (Dillon and Rigler 1974; Jones and Bachmann 1976) report models yielding higher predictions of chlorophyll and these matched our all lake results more closely. Both these studies reported slopes significantly  $> 1$ , which on a log–log plot produces an

**Table 4** Coefficient of determination ( $R^2$ ) and standardised regression ( $\beta$ ) coefficients for stepwise multiple regression model for total phosphorus and total nitrogen with chlorophyll in lakes split by alkalinity, depth and humic substances

Type	Clear			Humic			Polyhumic		
	$R^2$	Standardised coefficients		$R^2$	Standardised coefficients		$R^2$	Standardised coefficients	
		TP	TN		TP	TN		TP	TN
HS	0.46	0.47	0.28	0.34	0.59	–			
HVS	0.43	–	0.64	0.53	–	0.73			
MD	0.81	0.90	–	0.88	0.94	–			
MS	0.84	0.92	–	0.67	0.70	0.19			
MVS	0.81	0.90	–	0.73	0.63	0.28	0.63	–	0.80
LD	0.43	0.66	–	0.46	0.68	–			
LS	0.62	0.69	0.15	0.60	0.77	–	0.55	0.44	0.41
LVS	0.87	0.94	–	0.73	0.57	0.37	0.45	0.33	0.41

increasing yield of chlorophyll as TP increases; in contrast our all lakes model had a slope close to 1, showing that the relationship was close to linear. There are less reported models for total nitrogen, but those of Nürnberg (1996) and Prairie et al. (1989) are very similar to our all lakes combined model. However, our type-specific nutrient models produced values that correspond to the full range of those reported.

For  $\text{Chl} = f(\text{TP})$  our deep lakes model predicted the lowest chlorophyll concentrations, that are only slightly higher than those from the OECD equations. Our high-alkalinity lake model predicted higher values, while the models for the other lake types (low and moderate alkalinity shallow and very shallow lakes) predict chlorophyll values that are similar to the highest reported values. For  $\text{Chl} = f(\text{TN})$  our deep-lake equations provide lower predicted values than those in the literature, while the low and moderate alkalinity very shallow lakes are significantly higher. These discrepancies highlight some of the issues of log–log regression relationships: they hide within them a relatively wide range of values. Reynolds (1980) identified this when he pointed out that regression models can only describe the general behaviour of lake populations, and that in reality any particular lake will have a chlorophyll yield that is influenced by a number of factors other than nutrients. Nonetheless, these models are in many cases the only method for lake managers to determine the relationship between nutrients, a factor they can potentially control, and algal biomass. What is important is that managers should use an equation that is derived using data from a lake type that matches as closely as possible the lake they are concerned with.

Our data show that lake depth is a key factor, with deep lakes generally less responsive to nutrient enrichment. Stauffer (1991) pointed out the importance of self-shading and suggests this is responsible for the levelling of Chl at the highest TP concentrations. We might expect this in deep lakes if they are light limited with a euphotic depth less than the mixing depth; however, as we do not have data on these parameters we were unable to explore these relationships further. Shallow lakes, in particular very shallow lakes, are less likely to be light limited, particularly when chlorophyll values are low, as we found for low and moderate alkalinity lakes, and it is not surprising that in these lakes chlorophyll yield per unit of nutrient was generally higher.

We found the weakest relationships for  $\text{Chl} = f(\text{TP})$  in high alkalinity shallow lakes. Håkanson et al. (2005) also reported that the predictive power of Chl TP models increased markedly if hardwater lakes were omitted from their model and concluded that this was linked to bio-availability of particulate phosphorus. Calcium may have a role to play in precipitating phosphorus, making it less bio-available. High alkalinity lakes also generally have the highest levels of phosphorus which will have two effects that could influence the relationships: Firstly, chlorophyll will also be high, and even in shallow lakes could result in light inhibition, particularly for mixed or weakly stratified lakes. Secondly, at high TP the TN:TP ratio can be very low leading to nitrogen, rather than phosphorus limitation and thus a poor  $\text{Chl} = f(\text{TP})$  relationship. Prairie et al. (1989) concluded that variability in published relationships between Chl and nutrients can be accounted for by TN:TP ratios. We found similar results, but only for lakes where TN:TP was  $< 10$ . In these lakes, contrary to the findings of Prairie, we found that TN provided a better predictor of Chl than TP. A TN:TP ratio  $< 10$  is generally interpreted as evidence of nitrogen limitation (Downing and McCauley 1992; Smith 1982), so this observation is not surprising and may account for at least some of the scatter in the  $\text{Chl} = f(\text{TP})$  relationship in high alkalinity lakes.

For very shallow lakes, interactions with macrophytes are also likely to be important (Phillips 2003), and in lakes dominated by macrophytes, top-down control mediated through zooplankton grazing is also likely to play a key role in reducing chlorophyll concentrations below that predicted by their nutrient status (Scheffer 1990). We do not have sufficient information in our database to identify lakes likely to be influenced by top-down controls, but grazing by zooplankton is likely to introduce additional variability into the nutrient chlorophyll relationship and could be particularly strong in high alkalinity lakes where TP is within the range where top-down effects are strongest (Jeppesen et al. 2000). It is thus not surprising that we find the greatest variability in  $\text{Chl} = f(\text{TP})$  and we suggest that a combination of top-down influences, light inhibition and nitrogen limitation all contribute to limiting Chl in high alkalinity lakes.

Humic substances have also been shown to influence the nutrient chlorophyll relationships in lakes. Edmundson and Carlson (1998) and Havens (2003)

both reported lower yields of chlorophyll concentration per unit of TP. This may result from lower availability of light, due to increased water colour (Havens, 2003) or to lower bio-availability of phosphorus (Jones 1998). In our study we found that low and moderate alkalinity humic lakes had higher Chl concentrations. This was accompanied by higher TN and TP concentrations, but we could find no evidence of a reduced yield of chlorophyll. Nürnberg and Shaw (1999) also reported elevated nutrients in humic lakes, but in this case a small but significant positive effect on the  $\text{Chl} = f(\text{TP})$  relationship. We only found this for low alkalinity deep humic lakes, despite the fact that TN:TP ratios in humic low and moderate alkalinity lakes are lower than those in clear waters. Havens and Nürnberg (2004) pointed out the interactions between humic substances and mixing regime, demonstrating that the Chl concentration in mixed humic lakes is marginally reduced. They propose that in stratified lakes phytoplankton is able to adjust to reduced light availability by migration, resulting in elevated Chl level in surface layers. We can only speculate as to the explanation of the positive effect on the  $\text{Chl} = f(\text{TP})$  relationship we found in deep low alkalinity humic lakes, as we have insufficient information to identify stratified and mixed lakes. However, it is reasonable to assume that this lake type is very likely to be stratified, and the higher yield we observed may be a result of vertical migration of phytoplankton, an effect not seen in shallow mixed lakes. Also light adaptation in humic lakes yielding higher Chl:biomass ratios may contribute to explaining why deep low alkalinity humic lakes have steeper  $\text{Chl} = f(\text{TP})$  slope than the comparable clearwater lakes.

We found a clearer effect of TN:TP ratios on  $\text{Chl} = f(\text{TP})$ , particularly when combined with humic content. TN:TP was low in humic lakes and a substantial proportion of the very shallow humic lakes had values low enough to suggest TN rather than TP limitation. This is in contrast to Nürnberg and Shaw (1999), who found no evidence of changes to TN:TP ratios in humic lakes. In our dataset stepwise multiple regression demonstrated that TN rather than TP provided the best predictor of Chl in humic high alkalinity and polyhumic moderate alkalinity very shallow lakes and it seems likely that these lakes are more likely to be limited by nitrogen rather than phosphorus. We also conclude from our data that there is no significant effect of geographic region on

our regression models, despite the relatively wide area from which our data were drawn. A similar conclusion was reached by Seip et al. (2000), who concluded that trophic level, mediated via the TN:TP ratio, is more important in influencing the  $\text{Chl} = f(\text{TP})$  relationship than geographic region.

### Use for lake management

Regression models relating TP and TN to Chl are useful tools for lake managers as they enable management decisions to be made. For example, standards for nutrients can be established and actions taken within the catchment to achieve them. This is particularly important in the context of the Water Framework Directive where nutrient conditions capable of supporting good ecological status need to be defined. However, it is important to recognise the limitations of such models and consider the most appropriate way to use them.

Firstly, a linear model is only an approximation to what is more likely a curvilinear or sigmoidal relationship. The effects of this can be minimised by using type-specific regression relationships and restricting the regression to the linear regions of the response. In most cases  $\text{Chl} = f(\text{TP})$  provided the best model, but where TN:TP ratios are low ( $\leq 10$ )  $\text{Chl} = f(\text{TN})$  may be more reliable.

Secondly, it is important to be aware of the relative high degree of scatter hidden in log–log relationships. Many factors can influence this as we have discussed above and lake managers need to know how to deal with this uncertainty. For example, they may often need to assess permissible levels of nutrients necessary to prevent deterioration in ecological quality. A typical situation is the need to establish a target TP which would maintain a currently acceptable Chl concentration. By using an appropriate type-specific regression equation it would be possible to determine a TP concentration that would correspond to the desired Chl. However, there is significant uncertainty associated with this decision. It should be noted that Prairie (1996) cautions that regressions with  $R^2 < 0.65$  should not be used for predictions, as they have limited predictive power. Our combined lake model has an  $R^2$  which exceeds this value, but has the disadvantage that it does not distinguish the significantly different responses that lakes of different

types have. Although some of our type-specific models have  $R^2$  values  $> 0.65$  we suggest that it is more appropriate to use the models derived from grouped lake types where the effect of type on the regression equations within the group was not significant. Two of these models have  $R^2$  values  $> 0.65$ , but the high alkalinity lake model should be used with more caution.

While a prediction from a regression model provides the most likely outcome, for any particular lake, the outcome is uncertain with a chance that 50% of cases would yield higher or lower Chl concentrations than that predicted. In situations where it is important to reduce the risk of exceeding a given level of Chl, a target TP value could be determined using the upper confidence limit of the regression (e.g. 90th percent confidence of individual points). If TP targets were based on this relationship then on average only 10% of sites might exceed the desired Chl concentration, a more precautionary approach. In contrast, faced with setting a restoration target and justifying significant financial expenditure this might be interpreted as a too precautionary value. In this case, the regression line might be a more appropriate choice, as 50% of lakes are likely to achieve the desired Chl at the TP value predicted. To assist in this application to management, we provide figures containing both regression lines and confidence limits for our three type-specific models on a linear scale in Appendix 1, together with equations to predict both the regression line and values similar to the 90% confidence intervals based on the 5th and 95th percentiles of the residuals of the regression (Appendix 2).

## Conclusion

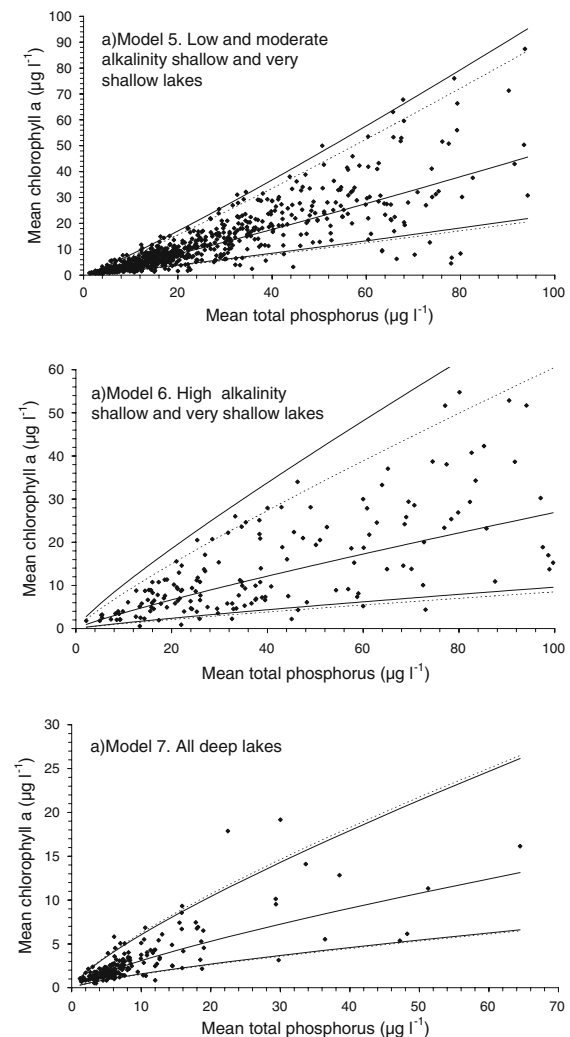
In conclusion we propose that chlorophyll nutrient regression models can provide useful tools for lake management, provided the uncertainty associated with them is clearly recognised. Total phosphorus and total nitrogen can be used to predict chlorophyll *a* concentration, although phosphorus is the better predictor for all types, with the exception of humic lakes.

We show that, except for deep lakes, the widely used OECD models predict much lower concentrations of chlorophyll than is predicted with our regressions, which are based on a large European

dataset. We thus present three models for a) deep lakes, b) low and moderate alkalinity shallow and very shallow lakes and c) high alkalinity shallow and very shallow lakes, which we suggest are the most appropriate for lake management in Europe, where more detailed regional models are not available.

## Appendices

**Appendix 1.** Relationship between summer mean total phosphorus ( $\mu\text{g l}^{-1}$ ) and chlorophyll *a* ( $\mu\text{g l}^{-1}$ ) in lakes grouped by alkalinity and mean depth. Solid lines show modelled regressions, upper and lower 90% confidence intervals. Dotted lines show modelled regressions  $\pm 95\text{th}$  percentile of regression residuals.



**Appendix 2.** Equations to calculate expected chlorophyll *a* from total phosphorus. The upper and lower boundary values are very similar to the 90% confidence limit of the regression and are determined by adding the 95th and 5th percentiles of the regression residuals. For any particular TP value approximately 90% of lakes will have a chlorophyll concentration below the upper boundary and above the lower boundary.

Model 5 Low and moderate alkalinity, shallow and very shallow lakes	
Upper boundary	$\text{Log}_{10}[\text{Chl}] = -0.528 + 1.108 \text{ Log}_{10}[\text{TP}] + 0.278$
Regression	$\text{Log}_{10}[\text{Chl}] = -0.528 + 1.108 \text{ Log}_{10}[\text{TP}]$
Lower boundary	$\text{Log}_{10}[\text{Chl}] = -0.528 + 1.108 \text{ Log}_{10}[\text{TP}] - 0.346$
Model 6 High alkalinity shallow and very shallow lakes	
Upper boundary	$\text{Log}_{10}[\text{Chl}] = -0.306 + 0.868 \text{ Log}_{10}[\text{TP}] + 0.352$
Regression	$\text{Log}_{10}[\text{Chl}] = -0.306 + 0.868 \text{ Log}_{10}[\text{TP}]$
Lower boundary	$\text{Log}_{10}[\text{Chl}] = -0.306 + 0.868 \text{ Log}_{10}[\text{TP}] - 0.500$
Model 7 All deep lakes	
Upper boundary	$\text{Log}_{10}[\text{Chl}] = -0.286 + 0.776 \text{ Log}_{10}[\text{TP}] + 0.306$
Regression	$\text{Log}_{10}[\text{Chl}] = -0.286 + 0.776 \text{ Log}_{10}[\text{TP}]$
Lower boundary	$\text{Log}_{10}[\text{Chl}] = -0.286 + 0.776 \text{ Log}_{10}[\text{TP}] + 0.305$

## References

- Canfield DE (1983) Prediction of chlorophyll *a* concentrations in Florida lakes: the importance of phosphorus and nitrogen. *Wat Res Bull* 19:255–262
- Clasen J, Bernhardt H. (1980) OECD Eutrophication programme. Shallow lakes and reservoirs project. Final report. Water Research Centre. Marlow UK.
- Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc* 74:829–836
- Dillon PJ, Rigler FH (1974) The phosphorus–chlorophyll relationship in lakes. *Limnol Oceanogr* 28:792–795
- Downing J, McCauley E (1992) The nitrogen–phosphorus relationship in lakes. *Limnol Oceanogr* 37:936–945
- Edmundson JA, Carlson SR (1998) Lake typology influences on the phosphorus–chlorophyll relationship in subarctic, Alaskan lakes. *Lake Reserv Manage* 14:440–450
- Field A (2005) *Discovering statistics using SPSS*. SAGE Publications, London
- Forsberg C, Ryding SO (1980) Eutrophication parameters and trophic state indices in 30 Swedish waste-receiving lakes. *Arch Hydrobiol* 89:189–207
- Håkanson L, Blenckner T, Bryhn C et al (2005) The influence of calcium on the chlorophyll–phosphorus relationship and lake secchi depths. *Hydrobiol* 537:111–123
- Havens KE (2003) Phosphorus–algal bloom relationships in large lakes of south Florida. *Lake Reserv Manage* 19:222–228
- Havens KE, Nürnberg GK (2004) The phosphorus–chlorophyll relationship in lakes: potential influences of colour and mixing regime. *Lake Reserv Manage* 20:188–196
- Jones C, Bachmann RW (1976) Prediction of phosphorus and chlorophyll levels in lakes. *J Water Pollut Control Fed* 48:2176–2182
- Jones RI (1998) Phytoplankton, primary production and nutrient cycling. In: Hessen DO, Tranvik L (eds) *Aquatic humic substances. ecology and biochemistry*. Ecological studies 133. Springer, Berlin, pp 145–175
- Moe SJ, Dudley B, Ptacnik R (2008) Databases: experiences from compilation and analyses of monitoring data from 5,000 lakes in 20 European countries. *Aquat Ecol*. doi: [10.1007/s10452-008-9190-y](https://doi.org/10.1007/s10452-008-9190-y)
- Myers R (1990) *Classical and modern regression with applications*. Duxbury, Boston, MA
- Nürnberg GK (1996) Trophic state of clear and colored, soft- and hard-water lakes with special consideration of nutrients, anoxia, phytoplankton and fish. *Lake Reserv Manage* 12:432–447
- Nürnberg GK, Shaw M (1999) Productivity of clear and humic lake: nutrient, phytoplankton, bacteria. *Hydrobiol* 382: 97–112
- OECD (1982) Eutrophication of waters monitoring assessment and control. Final report. OECD cooperative programme on monitoring of inland waters. OECD, Paris
- Phillips GL (2003) Eutrophication of shallow temperate lakes. In: O'Sullivan P, Reynolds C (eds) *The lakes handbook volume 2, lake restoration and rehabilitation*. Blackwell, Oxford, pp 261–278
- Prairie YT (1996) Evaluating the predictive power of regression models. *Can J Fish Aquat Sci* 53:490–492
- Prairie YT, Duarte CM, Kalff J (1989) Unifying nutrient–chlorophyll relationships in lakes. *Can J Fish Aquat Sci* 46:1176–1182
- Rast W, Lee GF (1978) Summary analysis of the North American (U.S.) OECD Eutrophication project: Nutrient load in-lake response relationships and trophic state indices. U.S. EPA-600/3-78-008
- Reynolds C (1980) Eutrophication and the management of planktonic algae: what Vollenweider couldn't tell us. In: Sutcliffe DW, Jones G (eds) *Eutrophication: research and application to water supply*. Freshwater Biological Association, Ambleside, pp 4–29
- Ryding SO, Rast W (1989) The control of eutrophication of lakes and reservoirs. UNESCO, Paris
- Sakamoto M (1966) Primary production by phytoplankton community in some Japanese lakes and its dependence on lake depth. *Arch Hydrobiol* 62:1–28

- Scheffer M (1990) Multiplicity of stable states in freshwater systems. *Hydrobiol* 200/201:475–486
- Seip KL, Jeppesen E, Jensen JP, Faafeng B (2000) Is trophic state or regional location the strongest determinant for Chl-*a*/TP relationships in lakes? *Aquat Sci* 62:195–204
- Smith VH (1982) The nitrogen and phosphorus dependence of algal biomass in lakes: an empirical and theoretical analysis. *Limnol Oceanogr* 23:1248–1255
- Stauffer RE (1991) Environmental factors influencing chlorophyll *v.* nutrient relationships in lakes. *Freshwat Biol* 25:279–295
- Van de Bund W, Cardoso AC, Heiskanen AS, Nöges P (2004) Overview of common intercalibration types. <http://wfd-reporting.jrc.cec.eu.int/Docs/typesmanual>
- Vollenweider RA (1976) Advances in defining critical loading levels for phosphorus in lake eutrophication. *Schweiz Z Hydrol* 37:53–84
- Vollenweider RA, Kerekes J (1980) The loading concept as a basis for controlling eutrophication philosophy and preliminary results of the OECD programme on eutrophication. *Prog Wat Rech* 12:5–38
- White E (1983) Lake eutrophication in New Zealand—a comparison with other countries of the Organisation for Economic Co-operation and Development. *N Z J Marine Freshwater Res* 17:437–444