 doi:10.1111/j.1365-2427.2007.01862.x

OPINION

A review and reassessment of lake phosphorus retention and the nutrient loading concept

MICHAEL T. BRETT AND MARK M. BENJAMIN
Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, U.S.A.

SUMMARY

1. We conducted a statistical reassessment of data previously reported in the lake total phosphorus (TP) input/output literature ($n = 305$) to determine which lake characteristics are most strongly associated with lake phosphorus concentration and retention. We tested five different hypotheses for predicting lake TP concentrations and phosphorus retention.
2. The Vollenweider phosphorus mass loading model can be expressed as: $TP_{out} = TP_{in}/(1 + \sigma\tau_w)$, where TP_{in} is the flow-weighted input TP concentration, τ_w is the lake hydraulic retention time and σ is a first-order rate constant for phosphorus loss.
3. The inflow-weighted TP input concentration is a moderately strong predictor ($r^2 = 0.71$) of lake phosphorus concentrations when using log–log transformed data. Lake TP retention is negatively correlated with lake hydraulic retention time ($r^2 = 0.35$).
4. Of the approaches tested, the best fit to observed data was obtained by estimating σ as an inverse function of the lake's hydraulic retention time. Although this mass balance approach explained 84% of the variability in log–log transformed data, the prediction error for individual lakes was quite high.
5. Estimating σ as the ratio of a putative particle settling velocity to the mean lake depth yielded poorer predictions of lake TP ($r^2 = 0.77$) than the approach described above, and in fact did not improve model performance compared with simply assuming that σ is a constant for all lakes.
6. Our results also demonstrate that changing the flow-weighted input concentration should always have a directly proportionate impact on lake phosphorus concentrations, provided the type of phosphorus loaded (e.g. dissolved or particulate) does not vary.

Keywords: hydraulic retention time, lake, mass balance, sedimentation, total phosphorus

Introduction

Lake water total phosphorus (TP) concentration is a strong predictor of numerous indicators of lake water quality and overall ecosystem composition (Peters, 1986), including: total phytoplankton biomass, the proportion of cyanobacteria and the prevalence of nuisance/toxic phytoplankton blooms (Jones & Bachmann, 1976; Downing, Watson & McCauley, 2001); water clarity and anoxia (Welch, 1992); the abundance, species diversity and species composition of macrophytes, zooplankton and fish (Jeppesen *et al.*, 2000) and the food quality of seston for herbivorous zooplankton (Müller-Navarra *et al.*, 2004).

Because of the strong associations between TP concentrations and lake ecology, numerous studies have attempted to determine the factors that exert the greatest impact on lake water TP concentrations. The pillar of this research has been the Vollenweider (1969,

Correspondence: Michael T. Brett, Department of Civil and Environmental Engineering, University of Washington, Box 352700, Seattle, WA 98195, U.S.A.
E-mail: mtbrett@u.washington.edu

 © 2007 The Authors, Journal compilation © 2007 Blackwell Publishing Ltd

1975, 1976) mass balance model for lake TP loading and retention, which provides a vehicle for predicting lake water TP concentrations based, in part, on lake morphometric and hydraulic data. Numerous studies have also attempted to develop general relationships between the morphometric/hydraulic characteristics of lakes and TP retention, either by extending Vollenweider's work or by developing independent models. The most frequently cited papers in this literature over the past 37 years are collated in Table 1.

Table 1 A summary of the main papers in the lake phosphorus mass balance literature. This table notes the year these papers were published, the number of citations they have received, the conceptual approach they used to characterize the basic mass balance equation and the parameter predicted in their analyses. These records were retrieved September, 2006, using the Web of Science®

Author	Formulation	Predicted parameter	Citations
Vollenweider (1969)	L, σ, z, q_s	Sed_{net}	183
Dillon & Rigler (1974)	L, R_P, z, ρ	TP_{lake}	191
Imboden (1974)	L, z, v, q_s	R_P	65
Chapra (1975)	R_P, v, q_s	R_P	41
Dillon (1975)	L, R_P, z, ρ	–	107
Dillon & Kirchner (1975)	R_P, σ, ρ	R_P	22
Kirchner & Dillon (1975)	L, R_P, z, ρ	R_P	75
Snodgrass & O'Melia (1975)	L, q_s, σ, z	TP_{lake}	41
Vollenweider (1975)	L, q_s, z, σ	R_P	341
Jones & Bachmann (1976)	L, z, ρ, σ	TP_{lake}	133
Larsen & Mercier (1976)	TP_{in}, R_P	R_P	97
Vollenweider (1976)	L, q_s, z	L_{CRIT}	499
Chapra (1977)	L_{QC}, Q, v, A_L	TP_{lake}	65
Ostrofsky (1978)	L_{QC}, Q, R_P	TP_{lake}	19
Schindler *et al.* (1978)	L, z, Q, V, T_w	TP_{lake}	105
Yeasted & Morel (1978)	$L, T_w, z, \sigma, v, R_P$	Lake trophic state	9
Chapra & Reckhow (1979)	TP_{in}, T_w	TP_{lake}	41
Reckhow & Chapra (1979)	L, R_P, q_s	R_P	16
Reckhow (1979)	L, z, T_w, q_s, v	TP_{lake}	17
Reckhow & Simpson (1980)	L, v, q_s	TP_{lake}	41
Uttormark & Hutchins (1980)	$TP_{in}, R_P, L, z, \rho, \sigma$	Lake trophic state	6
Canfield & Bachmann (1981)	L, R_P, z, ρ	TP_{lake}	106
Higgins & Kim (1981)	TP_{in}, σ, T_w	TP_{lake}	10
Nürnberg (1984)	L, R_P, q_s	R_P	84
Stauffer (1985)	L, q_s	TP_{lake}	33
Prairie (1988)	$L, R_P, q_s, \sigma, z, v$	Sed_{net}	9
Reckhow (1988)	TP_{in}, σ, T_w	TP_{lake}	17
Prairie (1989)	L, q_s, σ, z, R_P	Sed_{net}	7
Dillon & Molot (1996)	R_P, v, q_s	R_P	24
Nürnberg (1998)	L, R_P, q_s	TP_{lake}	9

The most highly cited studies in this literature have typically predicted lake TP concentrations based on input values for the areal phosphorus loading rate, mean lake depth, fractional phosphorus retention and areal hydraulic loading (see Table 2 for definitions of all terms and symbols used in this paper). Phosphorus losses within lakes are most commonly conceptualized as a function of the areal hydraulic loading rate or mean lake depth and a putative particle settling velocity. Studies have also modelled phosphorus losses as a function of the lake's hydraulic retention time. A smaller number of studies have suggested that the first-order rate constant for phosphorus loss is the same for all lakes or have applied empirical, multiple regression-based approaches to estimate expected phosphorus losses.

These studies have led to a widespread acceptance of the centrality of the mean depth and the areal TP loading rate as controlling parameters for in-lake TP concentrations, and of the ratio between a putative particle settling velocity and mean depth for the TP loss rate. We will argue that the approach typically

Table 2 Definitions of the terminology and symbols used for the lake TP mass balance calculations

A_L = surface area of the lake (m^2)
L = areal TP loading rate (mg TP m^{-2} $year^{-1}$)
L_{CRIT} = a critical areal TP loading rate above which some target lake TP concentration is likely to be exceeded (mg TP m^{-2} $year^{-1}$)
L_{QC} = conventional nutrient loading, i.e. $Q \times TP_{in}$ (kg $year^{-1}$)
Q = hydraulic inflow rate, usually quantified as an annual average (m^3 $year^{-1}$)
q_s = areal hydraulic loading rate, i.e. Q/A_L (m $year^{-1}$)
R_P = lake phosphorus retention (unitless)
$-r_{TP}$ = rate of TP loss from the lake (μg L^{-1} $year^{-1}$)
Sed_{net} = net phosphorus losses due to sedimentation, i.e. $L \times R_P$ (kg $year^{-1}$)
TP_{in} = flow-weighted TP influent concentration (mg m^{-3} or μg L^{-1})
$TP_{lake} = TP_{out}$ = TP concentration in the lake and its outflow (mg m^{-3} or μg L^{-1})
v = settling velocity of TP containing particles (m $year^{-1}$)
V_L = lake volume (m^3)
$\bar{z}$ = mean lake depth, i.e. V_L/A_L (m)
ρ = lake flushing rate, i.e. Q/V_L, or τ_w^{-1} ($year^{-1}$)
τ_w = mean hydraulic residence time, i.e. V_L/Q (years)
σ = first-order rate coefficient for TP loss from the lake ($year^{-1}$)

taken to conceptualize this problem has led to substantial confusion and an overestimate of the impact of these parameters on lake TP concentrations. We will present an alternative formulation for this problem that we believe makes a clearer link to the key processes regulating lake TP concentrations and retention. This approach has been suggested in previous studies, but it has not been documented as fully as in this paper, and it is not widely cited in the literature.

The Vollenweider equation

The Vollenweider (1976) phosphorus loading model is typically expressed mathematically as:

$$TP_{lake} = \frac{L}{\bar{z}(\rho + \sigma)}. \tag{1}$$

An important assumption in the derivation of eqn 1 is that the lake is well mixed and at steady state, so that the TP concentration throughout the lake and in the outlet stream can be characterized by a single value. The derivation also relies on the assumption that TP can be lost from the lake in only two ways: via advection (i.e. in the outlet stream) or via one or more first-order processes occurring within the lake. In the current context, the proportionality constant (or rate coefficient) is σ, i.e.:

$$-r_{TP} = \sigma(TP_{lake}). \tag{2}$$

In principle, σ could be a summation of rate coefficients for several first-order processes. However, σ is typically assumed to characterize net loss due to sedimentation of P-containing particles in the lake (Sed_{net}), as that is generally the dominant internal process affecting the TP concentration (Vollenweider, 1975). Nevertheless, it is recognized that this interpretation ignores several potentially important processes (Chapra, 1977), such as release of phosphorus from sediments, which may be important in lakes that experience extended periods of hypolimnetic anoxia, and which, when it occurs, is likely to not be a first-order process.

The areal loading rate of phosphorus, L, can be computed as follows:

$$L = \frac{Q(TP_{in})}{A_L}. \tag{3}$$

TP_{in} is calculated as the sum of all measured TP inputs to the lake (i.e. in tributaries, groundwater, surface precipitation and dry fallout), in mass/time, divided by the sum of all hydraulic inputs to the lake, in volume/time.

Finally, ρ can be expressed as:

$$\rho = \frac{Q}{V_L}. \tag{4}$$

The inverse of ρ (i.e. V_L/Q) is the mean hydraulic residence time, τ_w. Typically, ρ and τ_w are reported in inverse years ($year^{-1}$) and years (year) respectively.

Substituting eqns 3 and 4 in eqn 1, and noting that the mean depth, $\bar{z}$, equals V_L/A_L, we find:

$$TP_{lake} = \frac{Q(TP_{in})/A_L}{(V_L/A_L)((Q/V_L) + \sigma)}. \tag{5}$$

Multiplying the numerator and denominator by A_L/Q and cancelling like terms, we obtain:

$$TP_{lake} = \frac{TP_{in}}{1 + \sigma\tau_w}. \tag{6}$$

As has been pointed out by several authors (e.g. Vollenweider, 1975; Higgins & Kim, 1981; Welch, 1992), eqn 6 is identical to the classic result from chemical engineering for the relationship between the inlet and outlet concentrations of a substance that undergoes a first-order decay reaction in a continuous flow stirred tank reactor. While eqns 1 and 6 are mathematically equivalent, they cast the model in different terms, i.e. they identify different combinations of parameters as the key ones controlling TP_{lake}.

Lake TP retention

Mass balance models for TP behaviour in lakes, such as the Vollenweider model, are often expressed in terms of TP retention in the lake (R_P), defined as follows:

$$R_P \equiv \frac{TP_{in} - TP_{out}}{TP_{in}} = 1 - \frac{TP_{out}}{TP_{in}}. \tag{7}$$

If the lake is well mixed, TP_{out} can be equated with TP_{lake}. Furthermore, if the behaviour of the system is consistent with the Vollenweider model, then according to the model formulation shown in eqn 6, R_P can be computed as:

$$R_P = 1 - \frac{1}{1 + \sigma\tau_w} = \frac{\sigma\tau_w}{1 + \sigma\tau_w} = \frac{\sigma}{\rho + \sigma}. \quad (8)$$

Alternatives to the Vollenweider equation

Although the Vollenweider equation is the most widely used equation for predicting TP concentrations in lakes, a few other equations have been proposed (see Table 1). For instance, in a widely cited paper, Dillon & Rigler (1974) claimed to have developed an alternative approach for predicting TP_{lake}. When expressed in terms of the variables defined in this paper, their ultimate equation is as follows:

$$TP_{lake} = \frac{L(1 - R_P)}{\bar{z}\rho}. \quad (9)$$

However, because Dillon & Rigler used TP_{out} to calculate R_P, this equation can be transformed using the definitions for ρ, L and R_P accordingly:

$$TP_{lake} = \frac{(Q(TP_{in})/A_L)(1 - [1 - (TP_{out}/TP_{in})])}{(V_L/A_L)(Q/V_L)} = TP_{out}. \quad (10)$$

Thus, Dillon & Rigler's (1974) mass balance equation merely reiterates the basic assumption that the lake is well mixed.

Jones & Bachmann (1976) reported that an equation with the following form fit the data they analysed for 57 lakes:

$$TP_{lake} = TP_{out} = \frac{aL}{q_s + b\bar{z}}, \quad (11)$$

where a and b are adjustable parameters. Substitution of the expression QTP_{in}/A_L for L, and simplification, shows that eqn 11 can be written equivalently as:

$$TP_{lake} = \frac{a(TP_{in})}{1 + b\tau_w}. \quad (12)$$

In eqn 12, parameter a (which must have a value between 0 and 1) accounts for loss of a constant fraction $(1 - a)$ of the incoming TP, regardless of other lake properties; parameter b is a first-order rate constant for TP loss and is analogous to σ in eqns 1 and 6. Prairie (1988, 1989) developed a model that is mathematically equivalent to eqns 11 and 12.

Although several studies (e.g. Chapra, 1977; Reckhow & Chapra, 1983) have acknowledged that numerous factors can contribute to σ, by far the most common interpretation in the literature (see Table 1) is that σ is the ratio of the settling velocity (v) of P-containing particles to the mean lake depth, $\bar{z}$ (e.g. Chapra, 1975; Vollenweider, 1975; Larsen & Mercier, 1976), so eqn 6 can be rewritten as follows:

$$TP_{lake} = \frac{TP_{in}}{1 + (v\tau_w/\bar{z})} \quad (13)$$

When eqn 13 is used as the key model equation, v is treated as an adjustable parameter to provide the best fit to experimental data, whereas $\bar{z}$ is a known, lake-specific value.

The hypothesis that $\sigma = v/\bar{z}$ is consistent with a physical model in which the fraction of the TP that is associated with particles (and thus settleable) as well as the settling velocity distribution of those particles is constant throughout the year. This approach also implicitly assumes that the phosphorus associated with particles is permanently removed from the water column once the particles reach the sediments. However, both the particulate phosphorus fraction and the particle settling velocity distribution vary greatly during the course of the year in most lakes and especially between lakes, and phosphorus is often released from lake sediments. As a result, v is likely to have different values in different lakes, and during different seasons in a single lake, thereby reducing the predictive value of the equation.

As an alternative, it has been suggested that σ might be related to the inverse of the lake's hydraulic retention time via a function with the general form: $\sigma \propto \tau_w^x$, where x is an empirical constant. Larsen & Mercier (1976) suggested that such a correlation fits the available data reasonably well, with a value of x near −0.5.

Other researchers have proposed approaches for using the Vollenweider equation predictively in a way that does not focus on σ directly. That is, they have attempted to identify measurable parameters (usually related to lake morphology and hydrology) that correlate with R_P, without focusing on how those parameters relate to σ. By far the most common of these approaches focus on the areal hydraulic loading rate, the mean lake depth and a putative particle settling velocity as key variables (e.g. Kirchner & Dillon, 1975; Larsen & Mercier, 1976; Vollenweider, 1976).

While proposed relationships between TP_{lake} and L, $\bar{z}$, v and other parameters have generally not been framed in the context of hypotheses about σ, thinking of them in that way provides a rational and consistent way to compare the various hypotheses. In the remainder of this paper, we use both that approach and more conventional approaches to compare several of the hypotheses. In particular, we attempt to answer the question: To the extent that the overall Vollenweider model is correct, does one of the two ways of formulating the model presented earlier (in eqns 1 and 6) provide more insight into the parameters and processes that actually control TP concentrations in lakes? Or, equivalently, which parameters are most usefully thought of as master parameters for this application?

We address this question via an analysis of a database comprising 305 temperate and north temperate North American and European lakes. Previous studies that have attempted to assess the accuracy of the model equations have used smaller and often regional databases, and they have often selectively excluded many lakes from the analysis (see Dillon & Kirchner, 1975; Snodgrass & O'Melia, 1975; Larsen & Mercier, 1976; Vollenweider, 1976). We believe our analysis represents a more robust test of the validity of the model equations and the relative usefulness of the different formulations.

Hypotheses tested

The analysis was carried out by using several hypothesized relationships among key model parameters to predict TP_{lake} for each lake in the database. The hypothesized relationships are described in words below and mathematically in Table 5.

- Hypothesis 1: The rate coefficient σ is the same in all lakes, i.e. $\sigma = k_1$, where k_1 is a constant.
- Hypothesis 2: The rate coefficient σ is inversely proportional to the hydraulic retention time: $\sigma = k_2/\tau_w$, where k_2 is a constant.
- Hypothesis 3: The rate coefficient σ is related to a characteristic TP settling velocity (v) and the mean lake depth by: $\sigma = v/\bar{z}$. For the purposes of the current analysis, v is treated as an empirical constant that is the same in all lakes.
- Hypothesis 4: The rate coefficient σ is an inverse function of the lake's hydraulic retention time, i.e. $\sigma = k_4\tau_w^{x4}$, where k_4 and x_4 are constants.
- Hypothesis 5: The relationship between TP_{out} and other system parameters is given by an equation with the form proposed by Jones & Bachmann (1976) and Prairie (1988, 1989), i.e.

$$TP_{out} = \frac{aTP_{in}}{1 + b\tau_w}.$$

Methods

The sample

The data used in the analysis were compiled from previously published, large-scale lake TP input/output budgets. Lakes were included in this study if sufficient data were available to determine the TP loading to the lake, the in-lake TP concentration and the lake's basic morphometric and hydraulic characteristics (i.e. mean depth, surface area, volume and inflows). Among the lakes included in the analysis were the small, shallow, low hydraulic input, nutrient rich Iowa lakes investigated by Jones & Bachmann (1976); the large, very short hydraulic retention time TVA reservoirs analysed by Higgins & Kim (1981) and the mostly oligotrophic Canadian Shield forest lakes studied by Janus & Vollenweider (1981). Forty-four per cent of the lakes in the analysis were included in the USEPA (1975) national eutrophication survey of lakes in the north-east and north-central United States. When the same lake was reported in multiple studies, the most comprehensive data set was used. Fourteen lakes included in the source databases were excluded from this analysis because TP loading was indirectly estimated from land cover data. Eight other lakes were excluded because they had implausible input and output TP concentrations (e.g. $TP_{in} = 1430$ and $TP_{out} = 20$) or because the calculated input concentration could not be reconciled with the reported partial contributions from tributaries and point sources. Lakes were also excluded if they were suspected to be far from steady state (e.g. if they were recovering from nutrient enrichment; see Sas, 1989).

Statistical analyses

To test the five hypotheses we determined the parameter values that minimized the error sum-of-squares (ESS) between the predicted and observed TP_{lake} values using the Solver® (Microsoft Corporation, Redmond, Washington, U.S.A.) function in Excel® (Microsoft Corporation, Redmond, Washington, U.S.A.), for log–log transformed data. We also used Excel® to conduct sensitivity analyses examining how model performance varied with coefficient values. The optimal results for each hypothesis were confirmed using SPSS® (SPSS Inc., Chicago, Illnois, U.S.A.) v.14 for Windows®. A Bootstrap algorithm ($n = 30$) within SPSS® was also used to compute standard errors for the optimal coefficients. The results of all statistical calculations carried out in both Excel® and SPSS® (i.e. the r^2, ESS and optimal coefficient values) agreed to three significant figures.

The ESS and total sum of squares (TSS) were calculated as follows (Gallant, 1975):

$$\mathrm{ESS} = \sum[(\log(\mathrm{TP_{observed}}) - (\log(\mathrm{TP_{predicted}})]^2 = \sum\left(\log\frac{\mathrm{TP_{observed}}}{\mathrm{TP_{predicted}}}\right)^2, \tag{14}$$

and

$$\mathrm{TSS} = \sum\left(\log\frac{\mathrm{TP_{observed}}}{\mathrm{TP_{mean}}}\right)^2. \tag{15}$$

An r^2 value was then computed as follows:

$$r^2 = \frac{\mathrm{TSS} - \mathrm{ESS}}{\mathrm{TSS}}. \tag{16}$$

Because our statistical analyses were all based on log–log comparisons, the prediction interval (PI) in linear space for a TP mass balance model point prediction (x_0) is:

$$\mathrm{PI} = 10^{x_0 \pm \phi}, \tag{17}$$

where ϕ equals the prediction interval in log space, calculated as follows (Reckhow & Chapra, 1983):

$$\phi = t_{n-2}(\mathrm{SEE})\sqrt{1 + \frac{1}{N} + \frac{(x_0 - \bar{x})^2}{(N-1)S_x^2}} \tag{18}$$

where t_{n-2} is the appropriate two-tailed t-value for $n - 2$ degrees of freedom, SEE is the standard error of the estimate for each model, x_0 is the logarithm of the point prediction, $\bar{x}$ is the mean of all predicted log TP concentrations and S_x^2 is the variance of log $\mathrm{TP_{predicted}}$. The standard error of the estimate is calculated as:

$$\mathrm{SEE} = \sqrt{\frac{\mathrm{ESS}}{N - n_v}} \tag{19}$$

where N equals the sample size (i.e. 305) and n_v equals the number of variables in the models used to predict $\mathrm{TP_{lake}}$. As the prediction interval for an anti-log prediction is not symmetrical around the point prediction and the absolute value of the prediction interval is dependent on the point prediction, we also calculated a ratio between the upper and lower 95% prediction intervals, which we call the prediction interval ratio ($\mathrm{PIR_{95\%}}$), to provide a general measure of prediction error for the various models tested. As this ratio is slightly dependent on the point predictions, the ratios reported are always for the median observed TP concentration for this data set, which was 38 μg L^{-1}.

When retransforming $\log_{10}$-transformed model predictions back to linear space, a bias is introduced. This bias occurs because when a model is developed in log–log space its residuals are symmetrical in log space, but can be highly skewed towards larger values in linear space. This retransformation bias can be accounted for with the following correction factor (Sprugel, 1983):

$$\mathrm{CF} = 10^{2.303 * \mathrm{SEE}^2/2}. \tag{20}$$

Because both the prediction interval ratio and the retransformation bias correction factors are analytical solutions predicated on certain assumptions that may not be met for the types of log–log models we tested, we also calculated these values using a Bootstrap approach with replacement in Excel®. This was performed by first computing the predicted TP/observed TP ratio for each lake. These values were then copied 100 times and randomly sorted. A bootstrap mean ± 1 SE $\mathrm{PIR_{95\%}}$ for these data was derived by taking subsets of data with $n = 305$ and dividing the 97.5th percentile of the predicted TP/observed TP ratio by the corresponding value at the 2.5th percentile. The bootstrap CF-value was computed as the mean predicted TP/observed TP ratio from these simulations. Because these bootstrap estimates were already adjusted for sample size, the SE values for both the $\mathrm{PIR_{95\%}}$ and CF terms from these simulations were equivalent to the standard deviation values computed by Excel®.

Finally, we used the Akaike information criterion (AIC) (Akaike, 1970) to determine which of the models tested gave the best fit to the data while accounting for the fact that the models used different numbers of adjustable coefficients. The AIC statistic was calculated as:

$$\mathrm{AIC} = N \log\left(\frac{\mathrm{ESS}}{N}\right) + 2(n_v + 1), \quad (21)$$

where $n_v + 1$ represents the number of variables in the model plus the variance term in the AIC estimate. We then assessed these models according to the ΔAIC criterion (Burnham & Anderson, 2002), which is calculated as:

$$\Delta\mathrm{AIC}_i = \mathrm{AIC}_i - \min(\mathrm{AIC}) \quad (22)$$

where AIC_i represents the AIC value for model i and min(AIC) represents the lowest AIC value observed in the population of models assessed. According to Burnham & Anderson (2002), $\Delta\mathrm{AIC}_i$ values >10 suggest that model i is very unlikely to be valid, whereas values <2 suggest that the model is essentially equivalent to the optimal model. We then calculated Akaike weights as follows to determine the probability that model i is the best model amongst the population of models considered (Burnham & Anderson, 2002):

$$\text{Akaike weight} = \frac{e^{-\Delta\mathrm{AIC}_i/2}}{\sum_{\mathrm{all}\,j} e^{-\Delta\mathrm{AIC}_j/2}}. \quad (23)$$

The results of all these and other calculations are summarized in Table 5.

Our optimization approach is based on finding the best fit to TP_{lake} values (or, equivalently, TP_{out} values), as opposed to R_P values. The latter approach has been very commonly used (see Table 1). Our decision was based on the fact that TP_{lake} is strongly associated with many lake water quality and ecological characteristics. In contrast, TP retention has a direct effect on TP concentration, but only an indirect effect on other lake properties. In addition, as will be shown later, a small error when estimating R_P can manifest itself as a very large error in TP_{lake}, especially when R_P is high.

Results

Sample characteristics

An overview of the limnological characteristics of those lakes included in this analysis is provided in Fig. 1 and Table 3. Figure 2 shows that the lakes included in the analysis had a wide range of R_P values, with values at the 25th and 75th percentiles of 0.18 and 0.69 respectively. In 12% of the lakes, export of TP was reported to be greater than the TP input. Such a phenomenon could reflect release of TP from the sediments, or it could be due to errors in the estimates of TP_{out} and especially TP_{in}. For example, many lakes are sampled most intensively during the summer stratified period, when internal nutrient loading is most likely. During those times, many lakes have elevated water column phosphorus concentrations and net export of phosphorus, even if over the long term they trap a substantial portion of the phosphorus input (Welch, 1992).

A *t*-test using log-transformed data indicated that the lakes that appeared to have net phosphorus export did not have significantly different surface areas, volumes or input TP concentrations from those that had net TP retention. However, the lakes with net TP export did have somewhat smaller mean depths, higher in-lake TP concentrations, larger inflow rates and much shorter hydraulic retention times and much larger areal hydraulic loading rates than the others (Table 4).

The fact that net TP export was associated with short hydraulic retention time suggests sampling error might have been significant for these lakes. When the hydraulic retention time in a lake is short, TP_{in} and TP_{out} are often similar and errors in estimating TP_{in} could generate negative estimates for R_P. Nevertheless, we retained the lakes with negative R_P values in the analyses. This decision was based on the assumption that the types of errors leading to $R_P < 0$ probably permeate the entire data set, so that by excluding only cases with apparent net TP export, we would likely bias the data set towards lakes with positive errors for this value.

Hypotheses

Hypothesis 1 The hypothesis that the rate coefficient for TP loss (σ) is a constant led to reasonably good correspondence between predicted and observed TP values (Figs 3 & 4; Table 5), yielding a maximum r^2 of 0.77 for $\sigma = k_1 = 0.45 \pm 0.04$ (±1 SE) year^{-1}. This value of σ corresponds to a loss of approximately 0.12% of lake TP per day.

Hypothesis 2 Hypothesis 2 is that the product $\sigma\tau_w$ is constant, in which case TP_{out} is directly proportional to TP_{in} and R_P is the same in all lakes, independent of lake morphology and hydrology. As such, this hypothesis can be used to assess how much of the variability in lake TP concentrations can be explained

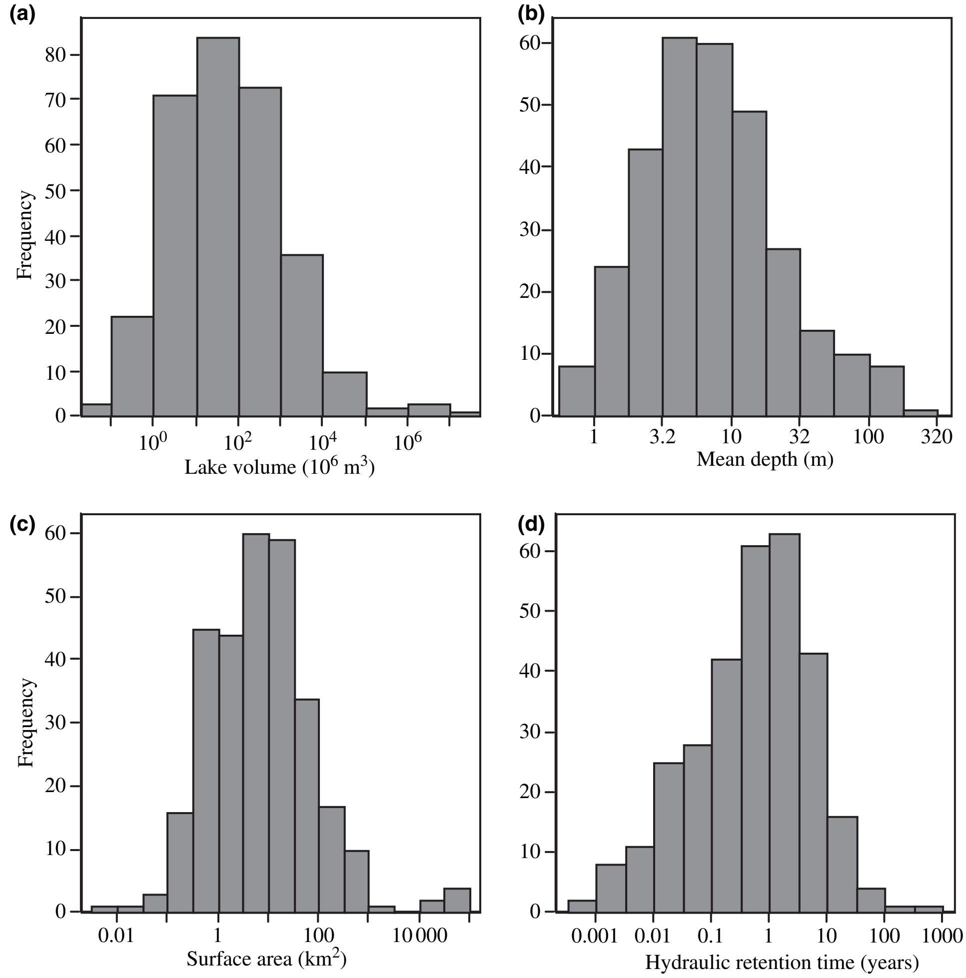

Fig. 1 The distribution of (a) lake volumes, (b) mean depths, (c) surface areas and (d) hydraulic retention times for the 305 lakes included in this analysis.

Table 3 General limnological characteristics of the lakes included in this analysis

Source	*N*	Percentile	Volume (10^6 m^3)	Surface area (km^2)	Mean depth (m)	Inflow rate (m^3 s^{-1})	Areal Hydr Loading (m $year^{-1}$)	Retention time (years)	TP_{in} (μg L^{-1})	TP_{out} (μg L^{-1})	TP_{out}/TP_{in} (unitless)
Overall	305	90th	2364	122	34.0	144	208	6.3	399	216	1.04
Overall	305	75th	346	26	14.3	18	57	2.6	194	92.0	0.82
Overall	305	50th	48.3	7.2	6.4	3.3	14	0.58	72.8	33.0	0.55
Overall	305	25th	4.42	1.2	3.2	0.48	4.0	0.10	31.4	12.1	0.31
Overall	305	10th	1.13	0.38	1.7	0.084	1.4	0.017	15.3	7.7	0.18
Vollenweider (1969)	8	50th	233	7.9	31.0	1.8	7.4	3.2	122	64.0	0.55
USEPA (1975)	134	50th	28.9	6.1	4.4	4.9	19.2	0.22	90.1	50.0	0.64
Jones & Bachmann (1976)	16	50th	5.22	3.5	2.0	0.25	1.33	1.6	259	86.5	0.34
Rast & Lee (1978)	30	50th	15.2	2.7	6.2	0.57	4.04	2.1	109	50.0	0.41
Fricker (1980)	20	50th	163	7.9	34.2	2.3	17.9	1.5	72.5	23.9	0.35
Ryding (1980)	14	50th	282	24	14.7	15	134	1.1	63.7	25.7	0.50
Higgins & Kim (1981)	18	50th	813	71	10.1	350	10.9	0.08	38.9	28.0	0.79
Janus & Vollenweider (1981)	65	50th	62.2	5.2	10.5	2.0	19.3	0.94	27.7	10.5	0.54

by considering the input TP concentration as the sole predictor. The best fit for this hypothesis was obtained for $k_2 = 1.06 \pm 0.08$ (Figs 3 & 4; Table 5), which yielded an r^2 value of 0.71 and a corresponding R_P value (equal to $(1 + \sigma\tau_w)^{-1}$) of 49%.

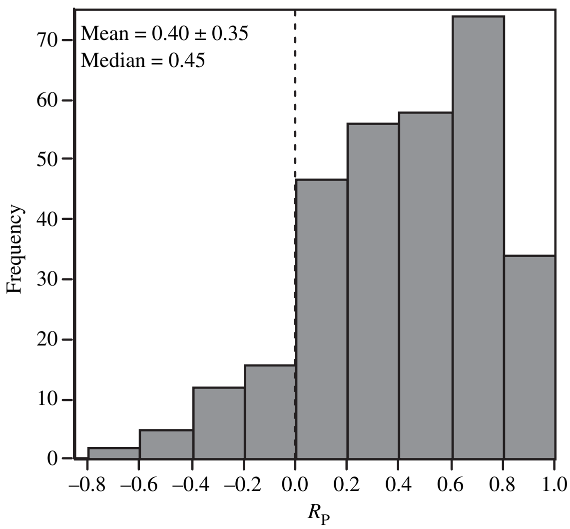

Fig. 2 A histogram of R_P values for the 305 lakes included in this study.

Hypothesis 3 The hypothesis that the rate coefficient for TP loss can be approximated as the ratio of a nominal TP settling rate (v) to the mean lake depth ($\bar{z}$) yielded a best-fit r^2 value of 0.77 when $v = 5.1 \pm 0.6$ m year^{-1} (Figs 3 & 4; Table 5).

Hypothesis 4 The hypothesis that the rate coefficient for TP loss can be represented by an expression with the form $\sigma = k_4\tau_w^{x_4}$ yielded an r^2 value of 0.84 for $k_4 = 1.12 \pm 0.08$ year$^{-0.47}$ and $x_4 = -0.53 \pm 0.03$ (Figs 3 & 4; Table 5). The variable x_4 has a value of −0.53 irrespective of the time units used (e.g. years, days or hours), while the magnitude of k_4 depends on the time units (e.g. k_4 equals 1.12 year$^{-0.47}$, 0.048 day$^{-0.47}$ and 0.0090 h$^{-0.47}$).

Hypothesis 5 Hypothesis 5 yielded r^2 values of 0.75 and 0.76 when the data set evaluated in this paper was tested using the values of *a* and *b* reported by Jones &

Table 4 Morphometric, hydrologic and nutrient characteristics of the lakes that retained ($n = 269$) and exported ($n = 36$) TP. The *P*-values are for *t*-tests carried out using log-transformed data

Parameter	Units	Positive R_P median	Negative R_P median	*t*-value	*P*-value
Surface area	km^2	7.2	5.8	0.57	0.5716
Volume	10^6 m^3	50	25	1.14	0.2544
Mean depth	m	6.9	5.1	1.96	0.0509
Hydraulic retention time	year	0.73	0.074	5.14	0.0001
Areal hydrologic loading	m year^{-1}	9.8	58	−5.14	0.0001
TP_{in}	µg L^{-1}	73.7	59.6	1.03	0.3032
TP_{out}	µg L^{-1}	30.0	70.5	−3.74	0.0002

Table 5 Summary of the hypotheses tested and optimum parameter values found. ESS is the error sum of squares, $PIR_{95\%}$ is the ratio between the upper and lower 95% prediction intervals, CF is the retransformation correction factor, n_v in the number of variables in each hypothesis, and AIC is the Akaike Information Criterion value. The error bars reported for the parameter values are ±1 SE

Hypothesis	Expression for TP_{lake}	Optimum parameter values	ESS	Optimal r^2	$PIR_{95\%}^{(a)}$	CF	n_v	AIC	ΔAIC	Ak. Wt.%
1	$\frac{Tp_{in}}{1+k_1\tau_w}$	$k_1 = \sigma = 0.45 \pm 0.04$ year^{-1}	22.6	0.767	11.8	1.22	1	−340.6	50	0.0
2	$\frac{TP_{in}}{1+k_2}$	$k_2 = 1.06 \pm 0.08$	28.5	0.706	15.9	1.28	1	−309.9	81	0.0
3	$\frac{TP_{in}}{1+\frac{v}{\bar{z}}\tau_w}$	$v = 5.1 \pm 0.6$ m year^{-1}	22.8	0.766	11.9	1.22	1	−339.8	51	0.0
4	$\frac{TP_{in}}{1+k_4\tau_w^{x_4}}$	$k_4 = 1.12 \pm 0.08$ year$^{-0.47}$ $x_4 = -0.53 \pm 0.03$	15.3	0.843	7.6	1.14	2	−390.7	0	99.2
5	$\frac{a(TP_{in})}{1+b\tau_w}$	$a = 0.65 \pm 0.03$ $b = 0.17 \pm 0.03$ year^{-1}	17.7	0.818	8.9	1.17	2	−371.2	19	0.8

(a) Defined as $\frac{\text{Upper } PI_{95\%}}{\text{Lower } PI_{95\%}}$.

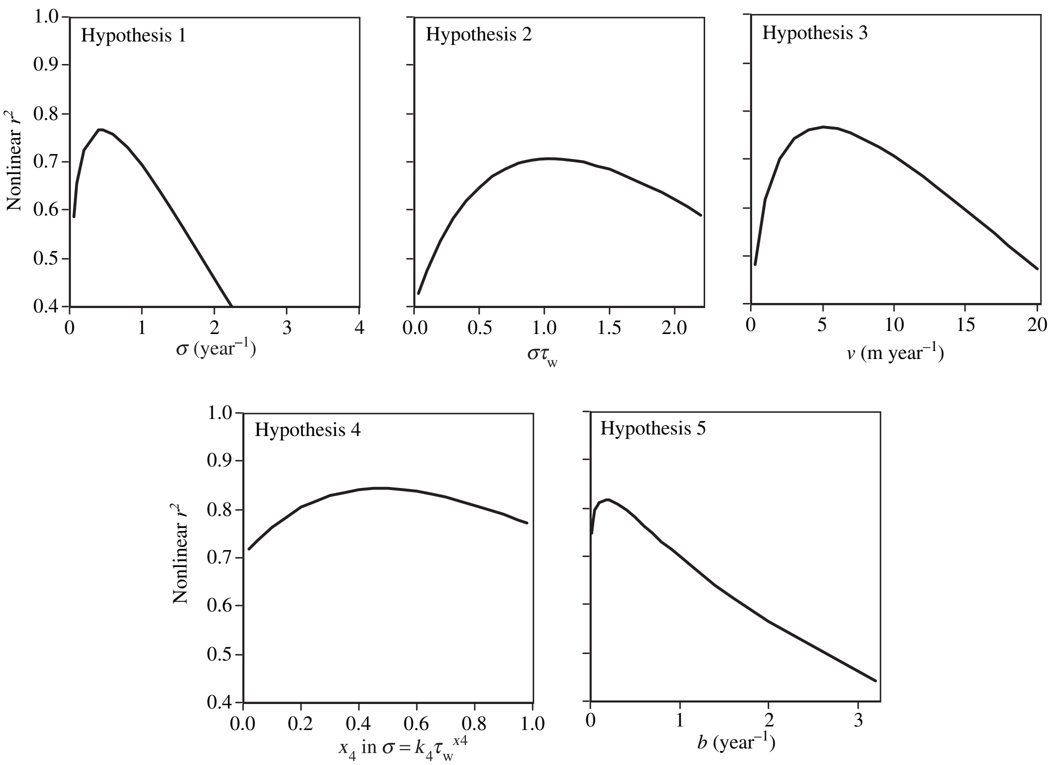

Fig. 3 Sensitivity analyses results for Hypotheses 1–5 for a range of plausible coefficient values. For Hypothesis 4, the optimal value of k_4 was determined for each value of x_4, and the r^2 value shown corresponds to that (x_4, k_4) combination. The same is true for Hypothesis 5, except that the parameters were b and a instead of x_4 and k_4.

Bachmann (1976) and Prairie (1988) respectively. When the values of a and b in eqn 12 were varied to obtain the best fit to the full data set evaluated in this paper, their values were 0.65 ± 0.03 and 0.17 ± 0.0.03 year^{-1}, respectively, and the r^2 value increased to 0.82 (Figs 3 & 4; Table 5).

Multiple regression models

For comparison, we also determined the best-fit values of the adjustable parameters for a multiple regression model for TP_{lake}, using the main morphometric and nutrient loading variables as inputs. The resulting best fit multiple regression model was:

$$TP = (0.71(\mu g/L)^{0.12}(yr)^{0.19}) * TP_{in}^{0.88} * \tau_w^{-0.19} \quad (24)$$

with an r^2 value of 0.84. Several alternative multiple regression models that included different sets of three, four or five morphometric and nutrient loading variables achieved fits to the original data that were only slightly weaker than $r^2 = 0.84$.

Based on the ΔAIC and Akaike weight criteria (see Table 5), Hypothesis 4 clearly outperformed the others. Furthermore, the ΔAIC values suggest the following overall ranking in model performance: Hypothesis 4 $\gg$ Hypothesis 5 $\gg$ Hypothesis 1 $\approx$ Hypothesis 3 $\gg$ Hypothesis 2. The fact that the Akaike weights strongly favoured Hypothesis 4, despite the fact that it only predicted 2.5% more variability than Hypothesis 5, was due to the large sample size.

The distribution of predicted TP/observed TP ratios for each hypothesis is depicted in Fig. 5. This figure shows the bootstrap $PIR_{95\%}$ and retransformation bias correction factor, each ±1 SE, for each of the hypotheses. This figure also shows that the bootstrap $PIR_{95\%}$ values for Hypotheses 1–5 were quite similar to the analytical $PIR_{95\%}$ values; in fact, all of the analytical $PIR_{95\%}$ values were within ±1 SE of the bootstrap means (see Table 5). Similarly, the analytical retransformation bias CFs were within ±1 SE of the bootstrap means for Hypotheses 4 and 5. However, the bootstrap retransformation bias CFs for Hypotheses 1–3 were

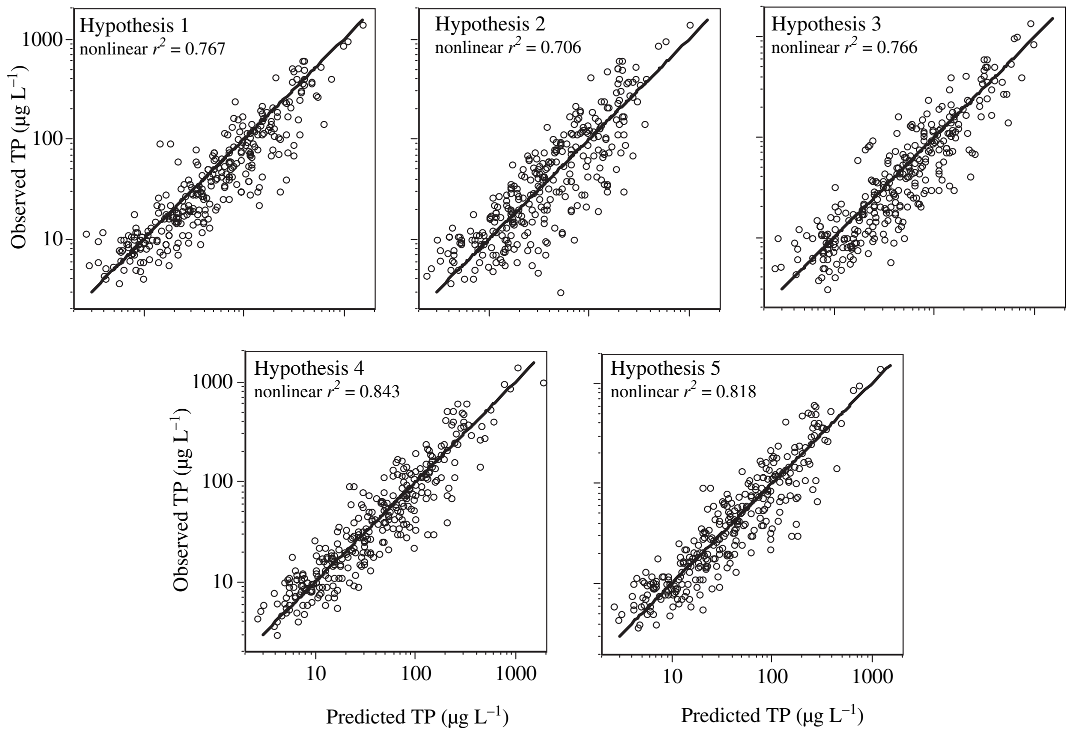

Fig. 4 Observed TP concentrations plotted against the optimal predicted results for Hypotheses 1–5 for the lakes used in this analysis. The diagonal line in these plots represents a perfect fit between predicted and observed values.

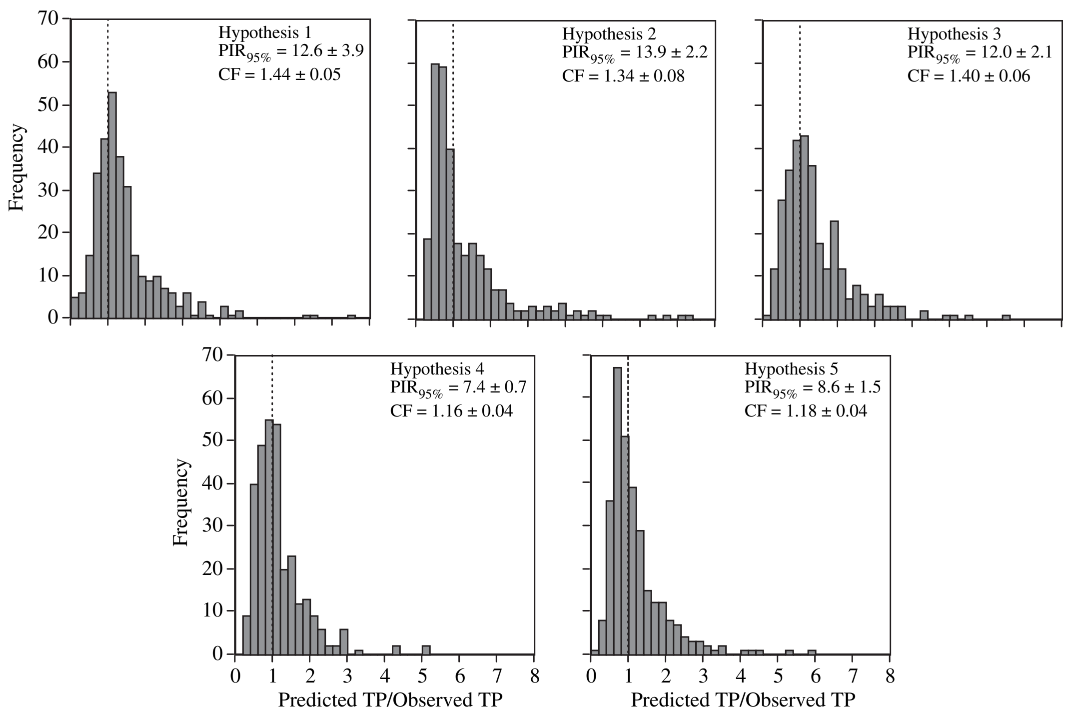

Fig. 5 The distributions of the Predicted TP to Observed TP ratio for each lake and each hypothesis assessed. The hatched vertical line represents the ideal 1 : 1 relation. The PIR$_{95\%}$ and retransformation bias correction factor (CF) values represent the mean ± 1 SE for Bootstrap simulations (n = 305). See the Methods section for a description of how these simulations were conducted and the respective values computed.

substantially higher than the analytical retransformation bias CFs for these hypotheses.

Discussion

These analyses show that the average input concentration (TP_{in}) was, by itself, a moderately strong predictor of lake TP concentrations. Simply knowing the flow-weighted input phosphorus concentration for this data set made it possible to statistically 'explain' 71% of the overall variability in the log-transformed lake TP concentrations. Schindler, Fee & Ruszczynski (1978) Yeasted & Morel (1978) and Reckhow (1988) also noted TP_{in} was the single best predictor of lake TP concentrations, but many other studies have failed to reach this conclusion. Incorporating the lake's hydraulic retention time in these predictions increased the explained variability to 77%. Finally, estimating the lake's TP concentration using both the lake's hydraulic retention time and a TP loss rate (σ) that was roughly proportional to $\tau_w^{-0.5}$ increased the variability explained to 84%. The most commonly applied approach for estimating TP losses in lakes, i.e. $\sigma = v/\bar{z}$, performed no better than simply assuming σ is a constant for all lakes. Despite a seemingly high r^2 value, the best overall model (i.e. Hypothesis 4) had substantial prediction error for individual lakes.

Comparison of the hypotheses based on predicted TP_{lake}

Hypothesis 4, and to a lesser extent Hypothesis 5, outperformed the other hypotheses in terms of providing the best fit between the predicted and measured TP_{lake} values. In contrast, a clear majority of previously published studies on this topic have used or advocated the particle settling velocity hypothesis (i.e. Hypothesis 3) (Chapra, 1975; Dillon & Kirchner, 1975; Kirchner & Dillon, 1975; Snodgrass & O'Melia, 1975; Vollenweider, 1976; Ostrofsky, 1978; Higgins & Kim, 1981; Nürnberg, 1984, 1998; Dillon & Molot, 1996). Our results show that this model construct led to much poorer predictions of TP_{lake} than Hypothesis 4, and no better predictions than the conceptually much simpler Hypothesis 1 (i.e. constant σ for all lakes). Furthermore, the best settling velocity found for Hypothesis 3, i.e. $v = 5.1 \pm 0.6$ m year^{-1}, is substantially lower than previously reported best-fit estimates, which are mostly in the range 8–16 m year^{-1} (e.g. Chapra, 1975; Vollenweider, 1975; Larsen & Mercier, 1976; Dillon & Molot, 1996). In contrast, phytoplankton cells and detrital material, which are the dominant particulates in most lakes, have settling velocities that are one (i.e. 55 ± 37 m year^{-1}) or two (i.e. 550 ± 73 m year^{-1}) orders of magnitude greater, respectively (see Burns & Rosa, 1980) than the best fit putative settling velocity for Hypothesis 3. Based on this evidence, we conclude that the available data do not support the widespread acceptance of the constant settling velocity model in the limnological literature.

We also found a variety of multiple regression models gave nearly identical results. This result is consistent with Reckhow & Chapra's (1983) observation that, because the terms typically considered in Vollenweider type analyses share many variables, various multiple regression models that superficially appear to be quite different yield very similar fits. This quandary is the principal reason why we opted not to use multiple regression as the primary tool in our analyses.

Our observation that Hypothesis 4 provided the best predictions for TP_{out} is consistent with Larsen & Mercier (1976), who reported the expression provided a good fit to a data set comprising 20 lakes with TP ≤ 25 μg L^{-1}. Although these authors are generally credited with first reporting that σ is proportional to $\tau_w^{-0.5}$, they did not advocate this approach over the particle settling approach, which gave nearly identical results in their analysis. Vollenweider (1976) also suggested that σ was proportional to $\tau_w^{-0.5}$, however his path to this result was much less clear. In fact, Vollenweider described his derivation of this result as being based on 'certain more or less defendable shortcuts' which yielded R_P values in the range 0.60–0.70 when $\tau_w = 1$ year and $R_P = 0$ when $\tau_w \rightarrow 0$.

The curious fact that the TP loss rate is approximately proportional to the inverse root of the lake hydraulic retention time has generated substantial discussion in the literature (e.g. Ahlgren, Frisk & Kamp-Nielsen, 1988; Welch, 1992). Schindler *et al.* (1978) and Canfield & Bachmann (1981) suggested that the reason that the TP loss rate increases when the hydraulic retention time decreases is that lakes with short hydraulic retention times receive relatively greater inputs of allochothonous, mineral-bound particulate phosphorus than do lakes with longer retention times. Because this particulate phosphorus is more susceptible to settling, short hydraulic retention time lakes, on average, have greater instantaneous TP loss rates than lakes with longer retention times. However, this explanation does not indicate why σ is best approximated as being proportional to the inverse square root of τ_w. We have explored this issue further using the current data set and found that the exponent value is dependent on which lakes are included in the analyses. For example, if the 20% of lakes with the longest hydraulic retention times are excluded, then σ is best approximated as nearly the inverse cube root of τ_w, i.e. $\sigma \approx \tau_w^{-0.35}$. This suggests that the proximity of x_4 to −0.5 when the whole data set is considered may be coincidental.

Comparison of the hypotheses based on predicted R_P

As has been noted, lake TP mass balance models have most often been evaluated based on their ability to predict observed R_P values. We therefore tested the accuracy of R_P predictions based on Hypotheses 1, 4 and 5, which provided the best fit to the TP_{lake} data. The tests for Hypotheses 1 and 4 were conducted by substituting the hypothesized relationship for σ into eqn 8 to obtain a predicted value for R_P, and comparing that predicted value to the observed value for each of the lakes in the database. Note that, according to eqn 8, R_P depends on both σ and ρ (or, equivalently, τ_w). As a result, the hypothesis that yields the best fit of σ to observed data will not necessarily yield the best fit to R_P. Hypothesis 5 was tested by rearranging eqn 12 to solve for the ratio TP_{out}/TP_{in} and then computing R_P as $1 - (TP_{out}/TP_{in})$ according to eqn 7.

The results of these analyses are shown in Fig. 6. All three models predict that R_P increases with increasing τ_w, as is observed empirically, especially for $\tau_w > 1$ month. This result is consistent with that reported by Yeasted & Morel (1978), who concluded that the lake hydraulic retention time had a stronger association with lake TP retention than did other lake morphometric characteristics. Figure 6 also shows that lakes with retention times of 1 month or less typically have TP_{lake} values that are similar to TP_{in} albeit with considerable scatter.

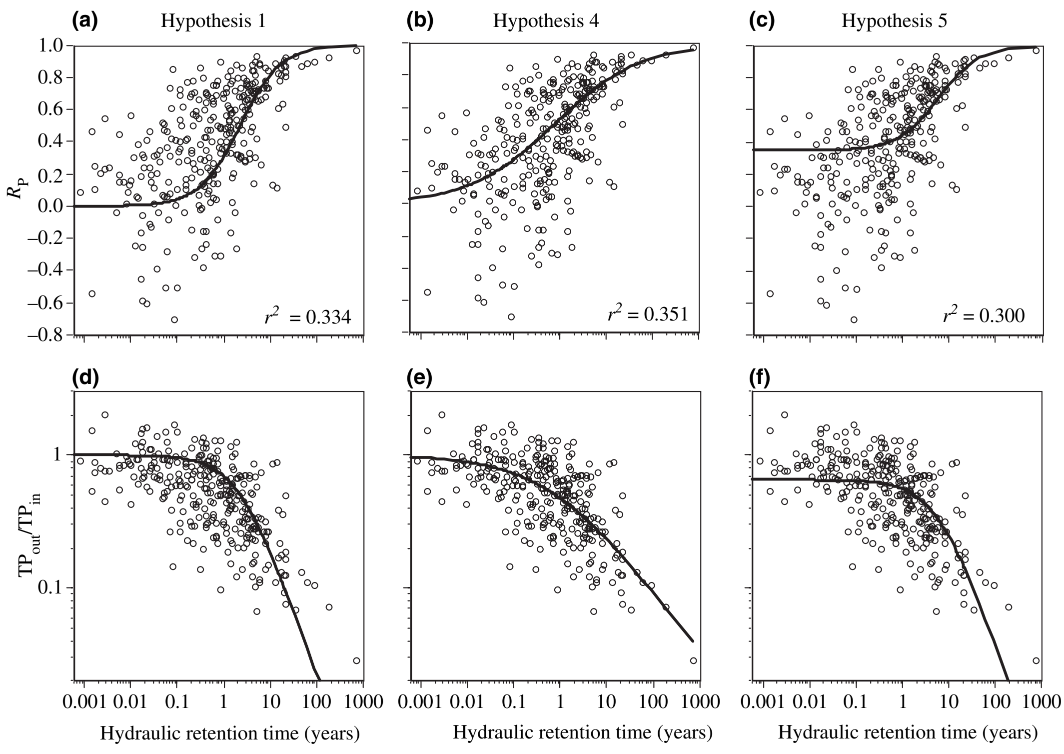

Fig. 6 A scatter plot of R_P (a–c) and the TP_{out}/TP_{in} ratio (d–f) versus lake hydraulic retention time for the 305 lakes considered in this analysis. The fitted curves reflect the predictions based on Hypotheses 1, 4 and 5 (see Table 5). The reported r^2 values are for linear models between observed and predicted R_P.

The R_P value predicted using the actual hydraulic retention time and a constant value of $\sigma = 0.45$ year^{-1} (based on Hypothesis 1) overestimated phosphorus losses to the sediments in 19 of the 22 lakes with $\tau_w > 10$ years. For this subset of lakes, the predicted R_P averaged 0.91, whereas the observed R_P value averaged 0.77, meaning the observed in-lake TP concentration was, on average, approximately 2.5 times the predicted value (23% of TP_{in} versus 9% of TP_{in}). Hypothesis 4 alleviated, but did not eliminate, these prediction errors (Fig. 6). The average R_P value predicted using this hypothesis for lakes with retention times greater than 10 years was 0.83. As a result, the actual TP_{lake} concentration was, on average, approximately twice the predicted value.

At the other end of the spectrum, Hypothesis 1 underestimated TP retention (i.e. it overestimated TP_{lake}) in most of the 71 lakes with retention times less than 1 month. In these lakes, the actual TP retention averaged 13%, compared with a predicted retention of only 1%. This is in part because this hypothesis is mathematically forced to $R_P \approx 0$ when $\tau_w \rightarrow 0$ year. Conversely, Hypothesis 4 predicted 15% TP retention in lakes with short hydraulic retention times, which was close to the observed average value. Thus, lakes with long retention times trap less TP than predicted by either Hypothesis 1 or 4, and lakes with short retention times trap more TP than Hypothesis 1 suggests.

We also tested whether the best-fit result we obtained for Hypothesis 4 was dependent on mean lake depth by separating our data set into shallow (i.e. ≤7 m, $n = 160$) and deep lakes (i.e. ≥10 m, $n = 109$) and refitting this hypothesis for these groups. For the subset of shallow lakes the best-fit net loss rate was as, $\sigma = 1.18\,\text{year}^{-0.48} * \tau_w^{-0.52}$ and for the deep lake subset this loss rate was $\sigma = 1.12\,\text{year}^{-0.46} * \tau_w^{-0.54}$. In terms of the phosphorus retention predicted by these alternative models, these equations are essentially indistinguishable from each other as well as from the overall best fit for Hypothesis 4.

The r^2 value for the correlation between predicted and observed R_P value was somewhat lower when the predictions were based on Hypothesis 5 than when they were based on either Hypothesis 1 or 4. The predictions for Hypothesis 5 were also somewhat better than those for Hypothesis 1 and somewhat worse than those for Hypothesis 4 for lakes with long hydraulic retention times. However, Hypothesis 5 provided very poor predictions for R_P in lakes with hydraulic retention times less than 1 month. In these lakes, the predicted TP retention averaged 35%, whereas the actual value was only 13%. Overall, Akaike weights for these three hypotheses suggest there was a 93% chance that Hypothesis 4 provided the best predictions for R_P, a 7% chance that Hypothesis 1 provided the best predictions and a 0% chance that Hypothesis 5 gave the best predictions.

Under many circumstances, optimizing the model based on the fit to either R_P or TP_{lake} data would yield similar results. However, when either the lake areal hydraulic loading rate is low or the hydraulic retention time is long, the two optimization approaches can yield dramatically different results. Under these conditions, a relatively small error in the estimated TP retention can manifest itself as a very large error in the predicted TP concentration. For example, Hypotheses 1 and 4 were almost equally good as predictors of R_P ($r^2 = 0.33$ and 0.35, respectively, see Fig. 6), but Hypothesis 4 was substantially better as a predictor of TP_{lake} (r^2 of 0.84, as opposed to 0.77 for Hypothesis 1). Interestingly, Hypothesis 5 was a good predictor of TP_{lake} but a much poorer predictor of R_P, particularly in lakes with short hydraulic retention times.

TP loading versus concentrations

Edmondson (1970) pointed out that in order to predict the impact of phosphorus loading on lake trophic nutrient concentrations, it is critically important to know whether a given TP load is associated with a large or small hydraulic load. The 'Cameron Lake paradox' (Dillon, 1975) provides an excellent example of this point. Dillon commented that 'although the phosphorus loadings were high, [Cameron] lake is not eutrophic because a high flushing rate counteracts the high loading'. In fact, as Dillon (1975) subsequently pointed out, TP loading in that lake was high because of the high flushing rate. As eqn 6 and Fig. 6 demonstrate, a high flushing rate does not counteract eutrophication – it exacerbates eutrophication, because it allows less time for whatever TP removal processes are active to occur. However, it should be noted that extreme flushing rates, i.e. ≥5% day^{-1} (when $\tau_w \leq 20$ days) can result in significant advective losses for phytoplankton and could result in lower phytoplankton biomass and especially less proportional representation by slower growing phytoplankton taxa such as cyanobacteria.

Cameron Lake was not eutrophic despite its very high areal phosphorus loading rate because the average input TP concentration to this lake was 16.5 μg L^{-1} (based on the values of L, Q and A_L reported in that paper). Thus, Cameron Lake received water which had TP concentrations below the 20 μg L^{-1} threshold for eutrophy, and TP retention in the lake further lowered TP_{lake} to 10.2 μg L^{-1}. Given the low input concentration, it was impossible for Cameron Lake to be eutrophic under steady-state conditions. Dillon (1975) acknowledged this point later when he stated 'it is easy to imagine the rapid passage of a large volume of water low in nutrients would yield a high loading rate but would not result in eutrophy'.

One can further demonstrate the relative importance of nutrient loading and input TP concentration by considering two scenarios that would both double the TP loading to a lake. In Scenario 1, the inflow doubles while TP_{in} remains constant, and in Scenario 2, TP_{in} doubles while inflow remains constant. For a lake with a hydraulic retention time of 0.6 year and a TP_{in} of 72.8 μg L^{-1} (the median values reported in Table 3), the TP_{lake} concentration predicted based on Hypothesis 4 is 38.7 μg L^{-1}. In Scenario 1, the expected value of TP_{lake} after the loading doubled would be 44.5 μg L^{-1}. However, in Scenario 2, the corresponding value would be 77.4 μg L^{-1}. In fact, according to any version of the Vollenweider mass balance model, doubling the inflow TP concentration will always cause the lake TP concentration to double.

In general, the relative impact of doubling inflows or doubling input concentrations on steady-state TP concentrations is dependent on lake retention time. According to Hypothesis 4, for the lakes in our analysis with hydraulic retention times less than 1 month, doubling the inflow rate would have only 4% as large an effect on TP_{lake} as would doubling TP_{in}. In lakes with hydraulic retention times greater than 10 years, doubling the inflow rate would have 31% of the effect on TP_{lake} as would doubling TP_{in}. These simple examples, and eqn 6, demonstrate that increasing input nutrient concentrations has a direct effect on lake trophic state, but increasing nutrient loading has only an indirect effect whose magnitude depends on lake hydraulic retention time and whether the increase in loading was due to changes in inflow amount or input concentrations. These results also suggest that if hydraulic loading to a lake is held constant, reducing phosphorus loading to a lake (i.e. reducing TP_{in}) should result in a proportional reduction in lake TP concentrations.

Model prediction error

This study has shown that a large proportion of the variability in lake TP concentrations can be explained by simply considering TP input concentrations and lake hydraulic retention time. The fact that Hypothesis 4 was able to explain 84% of the variability in log-transformed lake water TP concentrations may seem to suggest that this simple approach provides a clear description of the key factors influencing TP concentrations in lakes. However, it should be noted that the final r^2 value obtained is, in large part, simply a reflection of the extremely large range of lake types assessed vis-à-vis morphometric and hydraulic characteristics and especially TP_{in} concentrations. Hypothesis 4 explained only 35% of the variability in lake TP retention, and it had a Bootstrapped $PIR_{95\%}$ value of 7.4 ± 0.7 (±1 SE). This reinforces the idea that factors other than TP_{in} and τ_w play important regulatory roles for lake TP concentrations and retention, or that the underlying model assumptions of complete mixing, steady state and first-order losses are not valid for many lakes in the database.

The effect of incomplete mixing on phosphorus retention has been documented by Schindler *et al.* (1973) and Dillon & Molot (1996), who both noted that TP retention is strongly seasonally dependent; TP retention peaks during the summer stratified season when the epilimnion should be very well mixed and is at a minimum during the winter ice-covered period when only a very small fraction of the lake is well mixed. [These results can be partially reconciled by noting that the effective hydraulic retention time for temperate lakes will be much less during inverse thermal stratification.] Hypolimnetic anoxia can also result in seasonal phosphorus release from lake sediments (i.e. internal loading) (Nürnberg, 1984, 1998; Søndergaard, Jensen & Jeppesen, 2001; but see Gächter & Müller, 2003) that is not captured in the model construct. Weilenmann, O'Melia & Stumm (1989) have pointed out that natural coagulation processes have the potential to dramatically affect particle settling rates and therefore lake TP retention.

Other factors that can affect the efficiency of P trapping and that are also not included explicitly in the model include the amount of iron, calcium and/or aluminum in lake sediments (Boström, Jansson & Forsberg, 1982; Nürnberg, 1984); bioturbation by benthivorous invertebrates or fish (Boström *et al.*, 1982; Shapiro & Wright, 1984); and dominance of the zooplankton community by efficient grazers like *Daphnia* which can increase the downward flux of phosphorus-containing particulate matter incorporated in faecal pellets (Sarnelle, 1999).

Finally, the form in which phosphorus is supplied to the lake (i.e. inorganic or organic, dissolved or particulate) probably has a particularly important impact on phosphorus retention. If phosphorus is supplied predominantly as soluble reactive P or as a constituent of animal or human wastes, it can be readily incorporated into the biologic cycle of lakes (Reynolds & Davies, 2001). Conversely, if it is supplied predominantly in particulate form as P sorbed onto or incorporated into inorganic particles, it is much less bioavailable to phytoplankton or bacteria (Hatch, Reuter & Goldman, 1999; Pacini & Gächter, 1999; Reynolds & Davies, 2001). Furthermore, particulate, mineral-bound phosphorus will also settle out of the water column very rapidly and be lost to the sediments if the particles are large (i.e. sand sized) (Dorich, Nelson & Sommers, 1984; Hatch *et al.*, 1999; Pacini & Gächter, 1999). As Canfield & Bachmann (1981) noted for a database of 301 U.S. rivers, mean river TP and suspended sediment concentrations are clearly correlated ($r^2 = 0.45$), and river sediment transport varies greatly with local geologic conditions, land use and river gradient.

Conclusions

Lake TP concentrations are correlated moderately strongly with input TP concentrations, but only weakly with TP loading. Lake TP retention is inversely related to hydraulic retention time, but is not substantially correlated with mean lake depth or lake trophic status. Estimating the first-order rate coefficient for TP loss (σ) by an equation of the form $\sigma = k_4 \tau_w^{x_4}$ provides a better estimate of lake TP concentrations than several other relationships tested. Predictions based on estimating σ as the ratio of a nominal particle settling rate and mean lake depth (i.e. $\sigma = v/\bar{z}$) are no better than simply assuming σ was a constant for all lakes.

Lakes with long hydraulic retention times trap substantially less phosphorus than is predicted by models that assume that the rate coefficient is constant. Conversely, lakes with short hydraulic retention times trap substantially more TP than would be expected based on a model that does not account for the dependence of σ on τ_w. The best mass balance model tested in this study can explain 84% of the variability in log-transformed lake TP concentrations, but the same model explains only 35% of the variability in TP retention and has a large prediction error for individual lakes. The large prediction error is probably due to inadequate lake TP input budgets and the myriad of factors that are known to influence TP losses to lake sediments.

Acknowledgments

We thank Gertrud Nürnberg for helping us to compile the data set, and two anonymous reviewers for their constructive comments.

References

Ahlgren I., Frisk T. & Kamp-Nielsen L. (1988) Empirical and theoretical models of phosphorus loading, retention, and concentration vs. lake trophic state. *Hydrobiologia*, **170**, 285–303.

Akaike H. (1970) Statistical predictor identification. *Annals of the Institute for Statistical Mathematics*, **22**, 203–217.

Boström B., Jansson M. & Forsberg C. (1982) Phosphorus release from lake sediments. *Archiv für Hydrobiologie Beiheft Ergebnisse der Limnologie*, **18**, 5–59.

Burnham K.P. & Anderson D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer-Verlag, New York, NY.

Burns N.M. & Rosa F. (1980) In situ measurement of the settling velocity of organic-carbon particles and 10 species of phytoplankton. *Limnology and Oceanography*, **25**, 855–864.

Canfield D.E. Jr & Bachmann R.W. (1981) Prediction of total phosphorus concentrations, chlorophyll *a*, and Secchi depths in natural and artificial lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, **38**, 414–423.

Chapra S.C. (1975) Comment on an Empirical method of estimating the retention of phosphorus in lakes, by

W. B. Kirchner and P.J. Dillon. *Water Resources Research*, **11**, 1033–1034.

Chapra S.C. (1977) Total phosphorus model for Great Lakes. *Journal of the Environmental Engineering Division-ASCE*, **103**, 147–161.

Chapra S.C. & Reckhow K.H. (1979) Expressing the phosphorus loading concept in probabilistic terms. *Journal of the Fisheries Research Board of Canada*, **36**, 225–229.

Dillon P.J. (1975) The phosphorus budget of Cameron Lake, Ontario: the importance of flushing rate to the degree of eutrophy of lakes. *Limnology and Oceanography*, **20**, 28–39.

Dillon P.J. & Kirchner W.B. (1975) Reply to Chapra's Comment. *Water Resources Research*, **11**, 1035–1036.

Dillon P.J. & Molot L.A. (1996) Long-term phosphorus budgets and an examination of a steady-state mass balance model for central Ontario lakes. *Water Research*, **30**, 2273–2280.

Dillon P.J. & Rigler F.H. (1974) A test of a simple nutrient budget model predicting the phosphorus concentration in lake water. *Journal of the Fisheries Research Board of Canada*, **31**, 1771–1778.

Dorich R.A., Nelson D.W. & Sommers L.E. (1984) Algal availability of phosphorus in suspended stream sediments of varying particle size. *Journal of Environmental Quality*, **13**, 82–86.

Downing J.A., Watson S.B. & McCauley E. (2001) Predicting cyanobacteria dominance in lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 1905–1908.

Edmondson W.T. (1970) Book review (Water management research). *Limnology and Oceanography*, **15**, 169–170.

Fricker H. (1980) *OECD Eutrophication Programme Regional Project Alpine Lakes*. Swiss Federal Board for Environmental Protection & OECD. Dübendorf, Switzerland.

Gächter R. & Müller B. (2003) Why the phosphorus retention of lakes does not necessarily depend on the oxygen supply to their sediment surface. *Limnology and Oceanography*, **48**, 929–933.

Gallant A.R. (1975) Nonlinear regression. *The American Statistician*, **29**, 73–81.

Hatch L.K., Reuter J.E. & Goldman C.R. (1999) Relative importance of stream-borne particulate and dissolved phosphorus fractions to Lake Tahoe phytoplankton. *Canadian Journal of Fisheries and Aquatic Sciences*, **56**, 2331–2339.

Higgins J.M. & Kim B.R. (1981) Phosphorus retention models for Tennessee Valley Authority reservoirs. *Water Resources Research*, **17**, 571–576.

Imboden D.M. (1974) Phosphorus model of lake eutrophication. *Limnology and Oceanography*, **19**, 297–304.

Janus L.L. & Vollenweider R.A. (1981) *Summary Report, the OECD Cooperative Programme on Eutrophication Report, Canadian Contribution*. Canada Centre for Inland Waters, Burlington.

Jeppesen E., Jensen J.P., Søndergaard M., Lauridsen T. & Landkildehus F. (2000) Trophic structure, species richness and biodiversity in Danish lakes: changes along a phosphorus gradient. *Freshwater Biology*, **45**, 201–218.

Jones J.R. & Bachmann R.W. (1976) Prediction of phosphorus and chlorophyll levels in lakes. *Journal Water Pollution Control Federation*, **48**, 2176–2182.

Kirchner W.B. & Dillon P.J. (1975) Empirical method of estimating retention of phosphorus in lakes. *Water Resources Research*, **11**, 182–183.

Larsen D.P. & Mercier H.T. (1976) Phosphorus retention capacity of lakes. *Journal of the Fisheries Research Board of Canada*, **33**, 1742–1750.

Müller-Navarra D.C., Brett M.T., Park S., Chandra S., Ballantyne A.P., Zorita E. & Goldman C.R. (2004) Unsaturated fatty acid content in seston and trophodynamic coupling in lakes. *Nature*, **427**, 69–72.

Nürnberg G.K. (1984) The prediction of internal phosphorus load in lakes with anoxic hypolimnia. *Limnology and Oceanography*, **29**, 111–124.

Nürnberg G.K. (1998) Prediction of annual and seasonal phosphorus concentrations in stratified and polymictic lakes. *Limnology and Oceanography*, **43**, 1544–1552.

Ostrofsky M.L. (1978) Modification of phosphorus retention models for use with lakes with low areal water loading. *Journal of the Fisheries Research Board of Canada*, **35**, 1532–1536.

Pacini N. & Gächter R. (1999) Speciation of riverine particulate phosphorus during rain events. *Biogeochemistry*, **47**, 87–109.

Peters R.H. (1986) The role of prediction in limnology. *Limnology and Oceanography*, **31**, 1143–1159.

Prairie Y.T. (1988) A test of the sedimentation assumptions of phosphorus input–output models. *Archiv für Hydrobiologie*, **111**, 321–327.

Prairie Y.T. (1989) Statistical models for the estimation of net phosphorus sedimentation in lakes. *Aquatic Sciences*, **51**, 192–210.

Rast W. & Lee G.F. (1978) *Summary Analysis of the North American (U.S. Portion) OECD Eutrophication Programme: Nutrient Loading-Lake Response Relationships and Trophic State Indices*. Ecological Research Series, No. EPA-600/3-78-008, U.S. Environmental Protection Agency, Springfield, Virginia, U.S.A.

Reckhow K.H. (1979) Uncertainty analysis applied to Vollenweider phosphorus loading criterion. *Journal Water Pollution Control Federation*, **51**, 2123–2128.

Reckhow K.H. (1988) Empirical models for trophic state in southeastern U.S. lakes and reservoirs. *Water Resources Bulletin*, **24**, 723–734.

Reckhow K.H. & Chapra S.C. (1979) Error analysis for a phosphorus retention model. *Water Resources Research*, **15**, 1643–1646.

Reckhow K.H. & Chapra S.C. (1983) *Engineering Approaches for Lake Management, Vol. 1: Data Analysis and Empirical Modeling*. Butterworth Publishers, Boston, Massachusetts, U.S.A.

Reckhow K.H. & Simpson J.T. (1980) A procedure using modeling and error analysis for the prediction of lake phosphorus concentration from land-use information. *Canadian Journal of Fisheries and Aquatic Sciences*, **37**, 1439–1448.

Reynolds C.S. & Davies P.S. (2001) Sources and bioavailability of phosphorus fractions in freshwaters: a British perspective. *Biological Reviews of the Cambridge Philosophical Society*, **76**, 27–64.

Ryding S.O. (1980) *Monitoring of Inland Waters. OECD Eutrophication Programme. The Nordic Project*. Nordforsk, Helsinki.

Søndergaard M., Jensen J.P. & Jeppesen E. (2001) Retention and internal loading of phosphorus in shallow, Eutrophic Lakes. *The Scientific World*, **1**, 427–442.

Sarnelle O. (1999) Zooplankton effects on vertical particulate flux: testable models and experimental results. *Limnology and Oceanography*, **44**, 357–370.

Sas H. (1989) *Lake Restoration by Reduction of Nutrient Loading: Expectations, Experiences, Extrapolations*. Academia Verlag Richarz, St. Augustine, Germany.

Schindler D.W., Fee E.J. & Ruszczynski T. (1978) Phosphorus input and its consequences for phytoplankton standing crop and production in the Experimental Lakes area and in similar lakes. *Journal of the Fisheries Research Board of Canada*, **35**, 190–196.

Schindler D.W., Kling H., Schmidt R.V., Prokopowich J., Frost V.E., Reid R.A. & Capel M. (1973) Eutrophication of Lake 227 by addition of phosphate and nitrate: the second, third, and fourth years of enrichment, 1970, 1971, and 1972. *Journal of the Fisheries Research Board of Canada*, **30**, 1415–1440.

Shapiro J. & Wright D.I. (1984) Lake restoration by biomanipulation: Round Lake, Minnesota, the first two years. *Freshwater Biology*, **14**, 371–383.

Snodgrass W.J. & O'Melia C.R. (1975) Predictive model for phosphorus in lakes. *Environmental Science and Technology*, **9**, 937–944.

Sprugel D.G. (1983) Correcting for bias in log-transformed allometric equations. *Ecology*, **64**, 209–210.

Stauffer R.E. (1985) Relationships between phosphorus loading and trophic state in calcareous lakes of southeast Wisconsin. *Limnology and Oceanography*, **30**, 123–145.

USEPA (1975) *A Compendium of Lake and Reservoir Data Collected by the National Eutrophication Survey in the Northeast and Northcentral United States*. U.S. Environmental Protection Agency Working Paper No. 474, Corvallis, Oregon, U.S.A.

Uttormark P.D. & Hutchins M.L. (1980) Input/output models as decision aids for lake restoration. *Water Resources Bulletin*, **16**, 494–500.

Vollenweider R.A. (1969) Möglichkeiten und Grenzen elementarer Modelle der Stoffbilanz von Seen. *Archiv für Hydrobiologie*, **66**, 1–36.

Vollenweider R.A. (1975) Input-output models with special reference to the phosphorus loading concept in limnology. *Schweizerische Zeitschrift für Hydrologie*, **37**, 53–84.

Vollenweider R.A. (1976) Advances in defining critical loading levels for phosphorus in lake eutrophication. *Memorie dell'Istituto Italiano di Idrobiologia*, **33**, 53–83.

Weilenmann U., O'Melia C.R. & Stumm W. (1989) Particle-transport in lakes – models and measurements. *Limnology and Oceanography*, **34**, 1–18.

Welch E.B. (1992) *Ecological Effects of Wastewater: Applied Limnology and Pollutant Effects*. Chapman & Hall, London, U.K.

Yeasted J.G. & Morel F.M.M. (1978) Empirical insights into lake response to nutrient loadings, with application to models of phosphorus in lakes. *Environmental Science and Technology*, **12**, 195–201.

(Manuscript accepted 24 July 2007)

© 2007 The Authors, Journal compilation © 2007 Blackwell Publishing Ltd, *Freshwater Biology*, **53**, 194–211