

Analytical Methods for Parameter-Space Delimitation and Application to Shallow-Lake Phytoplankton-Dynamics Modeling

Gerrit van Straten

Twente University of Technology

P.O. Box 217

7500 AE Enschede, The Netherlands

ABSTRACT

The first step in parameter estimation is to reduce the dimensionality of the problem by deriving estimates from independent experimentation and from the literature. In addition, insensitive parameters are either removed or fixed. In the remaining lower-dimensional problem, parameter-space delimitation is possible by analytical means. Three conjunctive methods are derived: period-average analysis, extremum analysis, and quasisteady-state analysis. The basic idea is to find conditions for the parameters that must be fulfilled in order to comply with average and extreme values in the observations. The approach is applied to the modeling of the phytoplankton dynamics of Lake Balaton. The analytical techniques prove to supply valuable insight into parameter interrelationships and model adequacy, and can serve as satisfactory substitutes for formal parameter-estimation techniques in the early stages of model development.

PROBLEM DEFINITION

In lake eutrophication management the expected response of the phytoplankton levels in the lake to various management alternatives must be known. When phosphorus reduction is the most appropriate action to be taken, a first-order impression of its effectiveness can be obtained from a Vollenweider plot, which specifies a relationship between average summer chlorophyll-a and normalized phosphorus load [17,5]. Such a plot is based on observations on a large number of lakes, and the observed correlation can be viewed as a statistical model of lake response.

Although the statistical approach is attractive because of its simplicity, its drawbacks are also serious. First, the use of the generalized plot for an

individual lake is dangerous, because specific effects such as P exchange with sediments may cause a temporary deviation from the path predicted by the correlation model. Second, the Vollenweider plot is static; it cannot inform about the time at which algal peaks occur. Third, its use for evaluating other management alternatives such as flushing, dredging, and chemical dosage is limited or impossible. Finally, the uncertainty bounds are quite large, not the least because year-by-year differences in meteorological conditions are not explicitly accounted for. All these reasons explain why there has been and still is a strong interest in an approach which could overcome these difficulties: internally descriptive dynamic modeling.

There are numerous examples in the literature of internally descriptive models for phytoplankton dynamics. A review can be found in [9]. Generally, these models show quite a large number of parameters—at least ten and usually many more—while the data bases available for parameter estimation are quite limited. For example, in Lake Ontario, each month seven variables were measured in the epilimnion and six in the hypolimnion, which makes 160 spatially averaged data, whereas the well-known Manhattan model for Lake Ontario has 30–35 parameters [16]. In Lake Balaton, Hungary, four phosphorus fractions were directly measured fortnightly, which leads to a total of 100 values, while the simplest model, SIMBAL, has 26 parameters [15]. Consequently, for one year (which is the characteristic time unit because of the annual cycle), the data/parameter ratio is in the order of magnitude of 4 or 5. In addition, the total number of parameters is large, and these two aspects together make it very difficult to properly estimate the parameters by model calibration.

In this paper first a general procedure for reduction of the parameter-space dimension is discussed. Next, three analytical tools are offered for further delimitation of the regions of interest within the parameter space, as an intermediate step prior to parameter estimation by formal calibration routines. These tools are extremum analysis, period-average analysis, and quasi-steady-state analysis. The basic idea is to derive conditions for relationships between parameters which must be obeyed approximately in order to achieve a model behavior that matches the period-average, extremum, and (quasi-)steady-state observations in the system. This novel approach is then applied to a phytoplankton dynamics model for the Keszthely Basin, the most polluted, south-western part of Lake Balaton. The usefulness of the method is illustrated by showing the goodness of fit based on the parameter sets obtained from the analysis without further additional calibration. Even more important is the additional insight into parameter sensitivity and interconnectedness, into the role of information shortage and data uncertainty, and into model adequacy.

THEORY

Strategy of Reducing the Number of Calibration Parameters

The modeler, having finished the structural outline of his model, is next confronted with the task of estimating p parameters by comparing the model results with actual observations. Usually, this is done by defining some sort of objective function, which measures the difference between the model output and behavior of the system, followed by a procedure to minimize the value of the objective function. Numerous automatic minimum seeking routines have been proposed in the literature (see e.g. [2]) and have been implemented in numerical libraries. However, when p is large (over 5–10, say), it turns out in practice that useful results are difficult to obtain. One of the reasons is that the objective-function hypersurface in the $(p + 1)$ -dimensional space may not be smooth, and may show several local minima. In addition, with unconstrained optimization, parameters often dwell off to nonfeasible regions. Both phenomena occur more heavily in problems with more than one state variable, because the values of the objective function for a good fit of variable A and a less good fit of variable B , and vice versa, can be comparable, whereas the associated parameter sets may be entirely different. So there is a need to take fixed values for as many parameters as possible. A crucial step in a parameter-estimation effort in complex models is therefore to decide which parameters should be fixed and which should be calibrated. Before discussing this decision in more detail, it is appropriate to issue a warning. From a probabilistic point of view, fixing a parameter value is equivalent to the assumption that this parameter is exactly known. As a consequence the estimation error for the remaining parameters is reduced [1]. Furthermore, Figure 1 shows that the best-fit values of calibration parameters under the condition of others being exactly known may be quite different from the true best-fit values. This must be recognized if one wishes to attach a physical meaning to the calibration parameters: the values found, being conditional, usually have no general significance. But if making predictions is the major concern, the parameter bias is of less importance, because, after all, the conditional parameter set still led to an acceptable fit.

For deciding which parameters should be fixed, two attributes are of interest. The first derives from the question whether confident information is available or can be obtained on the parameter value. If such information is available, it should be used. Sources of information are:

- (i) the literature, and
- (ii) independent experimentation.

A number of phytoplankton dynamics parameter values have been published;

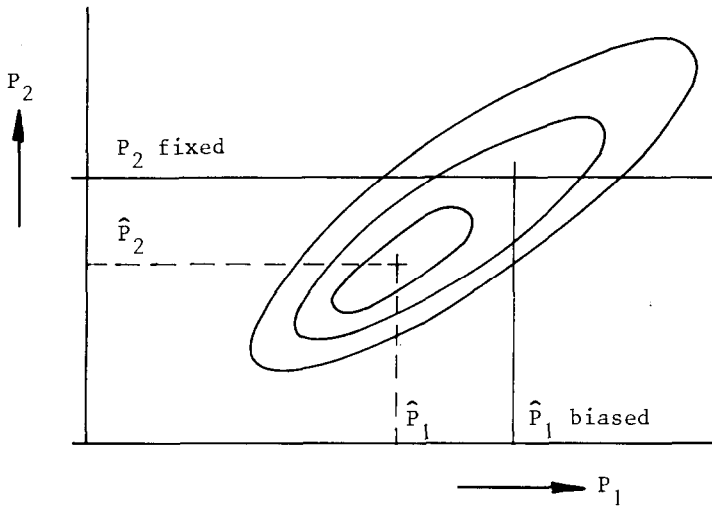


FIG. 1. Fixation of parameters leads to bias in the remaining calibration parameter estimates.

see e.g. [6]. However, some skepticism is advisable, because many parameters are model bound (and thus not transferable to other models) due to the various types of model simplifications applied. Examples are spatial averaging (affecting dispersion and mixing coefficients), integration of variables (whole-community algal parameters versus single-species parameters), aggregation of processes (simple first-order mineralization assumption covering a whole world of underlying chemical and microbial processes), etc. Perhaps the best source of information is independent experimentation. The basic principle is to isolate processes from the system and study them separately either in situ or in the laboratory. Examples are the estimation of algal-growth and light-limitation parameters from vertical primary production profiles, the measurement of sediment phosphorus-exchange parameters in undisturbed sediment columns, the oxygen uptake rates in sediment caps, etc. Again, of course, there are some questions on correspondence between the isolated and the full system [7], but with careful design good results are possible.

The second attribute is sensitivity. If a sensitivity analysis reveals that the model output is insensitive to a parameter, then there is no chance to estimate that parameter from curve fitting. Now, if confident information is available on that parameter, its value can be fixed. The occurrence of the parameter in the model may enhance the model's predictive power if future management drives the system to another working point—that is, the model may not be sensitive to the parameter now, but it may become sensitive under changed

TABLE 1
STRATEGY FOR SELECTING CALIBRATION PARAMETERS

	Confident	Nonconfident
Sensitive	Fix or calibrate within narrow limits	Calibrate
Nonsensitive	Fix; occurrence in the model may enhance the model's predictive power	Occurrence may be questioned. If the model cannot be restructured, fix.

management. On the other hand, for insensitive parameters for which no confident information can be obtained the situation is more complex. One might question whether it is justified for such parameters to occur at all. Restructuring of the model by further simplification must be seriously considered. If this is not possible, the most appropriate strategy is to assume some fixed value, without putting much effort into it, because the model is insensitive to it anyhow.

The strategy outlined above is summarized in Table 1. It amounts to the statement that calibration parameters should be those for which the model is sensitive, and for which no confident information from other sources is available.

Reduction of Parameter Space by Analytical Techniques

The procedure outlined above leads to a reduction of the dimensionality of the problem, from the original p -dimensional parameterspace to a n -dimensional space, where n is the number of parameters left for calibration, and $n < p$. The next step is to restrict the parameter space itself by looking for interrelationships between parameters, thus fencing in feasible regions. Let the model be formulated as

$$\dot{x} = f(x, u, p_f, p_c), \tag{1}$$

where

- $x(t)$ is the vector of state variables (dimension s),
- $u(t)$ is the input vector (dimension m),
- p_f is the vector of fixed parameters (dimension $p - n$),
- p_c is the vector of calibration parameters (dimension n).

Let

$$\mathbf{y} = H\mathbf{x} \quad (2)$$

be the observations ($\dim q \leq s$), and H a linear transformation matrix. Now define three types of processed observation data:

- (i) $y_i(t^*)$, the extreme of observation variable y_i at t^* ,
- (ii) \bar{y}_i , the average of y_i over a period $t_1 - t_2$,
- (iii) \tilde{y} , the steady-state value of y at constant inputs $\tilde{\mathbf{u}}$.

Items (i) and (ii) can be read directly or calculated from the sequence of observations. This is not the case with the steady-state value \tilde{y} , because the inputs in environmental systems will almost never be truly constant for a period of time sufficiently long to allow steady state to be reached. There are, however, two conditions for which quasisteady state can be assumed, allowing the computation of \tilde{y} :

(a) If y follows the variations in \mathbf{u} rapidly, steady state is reached "immediately" all the time. The dynamic behavior is more or less a sequence of steady states.

(b) If y follows the variations in \mathbf{u} slowly, then the inputs may be averaged over a period where y is approximately constant, and the response is a quasisteady state from which the fluctuations are filtered by the system.

With the three variables defined, three types of analysis are possible from which relationships between parameters derive that must be fulfilled to comply with the actual data: extremum analysis, period-average analysis, and quasisteady-state analysis.

Extremum Analysis. At the extremum of y_i we have

$$\dot{y}_i(t^*) = 0; \quad (3)$$

hence, with Equations (1) and (2),

$$\mathbf{h}_i^T \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*, \mathbf{p}_f, \mathbf{p}_c) = 0, \quad (4)$$

where \mathbf{h}_i^T is row i of the matrix H , and $\mathbf{x}^*, \mathbf{u}^*$ are shorthand for $\mathbf{x}(t^*), \mathbf{u}(t^*)$, respectively. Equation (4) specifies a condition for the choice of the parameters \mathbf{p}_c which must be fulfilled at an extremum for the variable y_i . In order to apply this equation the values of \mathbf{u}^* and \mathbf{x}^* must be known. For the inputs \mathbf{u}^* this is no problem, because the times of the extrema are known from the

observations and hence \mathbf{u}^* can be read from graphs or function tables. However, the situation for \mathbf{x}^* is different, because \mathbf{x}^* must be reconstructed from the observations. This is possible by writing

$$\mathbf{x}^* = H^{-1}\mathbf{y}^*, \tag{5}$$

but only if H is a square matrix, i.e. $q = s$, and nonsingular. In cases where the number of observation variables exceeds the number of state variables (i.e. $q > s$), $q - s$ observation variables are redundant, and the problem can be cast in a form where $q = s$. Of course, Equation (5) is only a rough state observer because the measurements are not exact. For the purpose of parameter-space delimitation this is not a serious problem.

Note that, in extremum analysis, at the time of an extremum for \mathbf{y}_i there is no need for the other \mathbf{y} -elements to be at an extremum as well.

Period-Average Analysis. Integration of \mathbf{y}_i over the period t_1 to t_2 leads to

$$\frac{1}{\Delta t} \int_{t_1}^{t_2} \dot{\mathbf{y}}_i(t) dt = \frac{1}{\Delta t} \mathbf{h}_i^T \int_{t_1}^{t_2} \dot{\mathbf{x}}_i(t) dt = \frac{1}{\Delta t} \mathbf{h}_i^T \int_{t_1}^{t_2} \mathbf{f} dt,$$

i.e.,

$$\frac{\Delta \mathbf{y}_i}{\Delta t} = \mathbf{h}_i^T \overline{\mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p})} = \sum_{j=1}^s h_{ij} \overline{f_j(\mathbf{x}, \mathbf{u}, \mathbf{p})}, \tag{6}$$

where $\Delta \mathbf{y} = \mathbf{y}_i(t_2) - \mathbf{y}_i(t_1)$, $\Delta t = t_2 - t_1$, and the overbar indicates time averaging.

A problem in the application of (6) is that in general $\overline{\mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p})} \neq \mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{p}})$, and so the computation of $\bar{\mathbf{f}}$ is not straightforward. A special case arises if the right-hand side of (6) is linear in the state variables and if the inputs do not occur in the coefficients, i.e. if it is possible to write

$$\sum_{j=1}^s h_{ij} f_j(\mathbf{x}, \mathbf{u}, \mathbf{p}) = \mathbf{a}^T \mathbf{x} + \mathbf{b}^T \mathbf{u}, \tag{7}$$

where the coefficient vectors only depend on the parameters \mathbf{p} . In this special case

$$\sum_{j=1}^s h_{ij} \overline{f_j(\mathbf{x}, \mathbf{u}, \mathbf{p})} = \mathbf{a}^T \bar{\mathbf{x}} + \mathbf{b}^T \bar{\mathbf{u}}, \tag{8}$$

and (6) becomes

$$\frac{\Delta y_i}{\Delta t} = \mathbf{a}^T H^{-1} \tilde{\mathbf{y}} + \mathbf{b}^T \tilde{\mathbf{u}}, \tag{9}$$

where $\bar{\mathbf{x}}$ has been replaced by $H^{-1} \tilde{\mathbf{y}}$ under the assumption that H^{-1} exists.

Quasisteady-State Analysis. At steady state $\dot{\mathbf{y}} = \mathbf{0}$ and so $\dot{\mathbf{x}} = \mathbf{0}$. Consequently, the (quasi)-steady-state parameter relations are given by

$$\mathbf{f}(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}, \mathbf{p}_f, \mathbf{p}_c) = \mathbf{0}, \tag{10}$$

where $\tilde{\mathbf{x}} = H^{-1} \tilde{\mathbf{y}}$, and $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{u}}$ are instantaneous values for rapid processes, or averaged values over time for very slow processes, as discussed previously. Again, H^{-1} must exist, and the same remarks apply as before.

APPLICATION

Lake Balaton

The approach is applied to the modeling of phytoplankton dynamics in the shallow Lake Balaton, Hungary, the largest lake in central Europe. We confine the analysis to the Keszthely Basin (see Figure 2), the most polluted bay of the lake. The major tributary, the Zala River, enters the lake in the Keszthely region, draining an area of 2622 km². The observations used are chlorophyll-a, dissolved reactive phosphorus, and total phosphorus, and are shown in Figure 3 for the year 1977. The inputs to the system are global radiation (R), temperature (T), dissolved inorganic P load (L_P), and particulate P load (L_D), shown in Figure 4. Some relevant data are collected in Table 2. Management interest is concentrated on methods to reduce the phytoplankton biomass, and extensive research has been conducted, including modeling, to assist in the design of proper management options [11, 12].

Phytoplankton Dynamics Model

In this analysis we adopt a modification of the simple phytoplankton model SIMBAL [14]. The structure of the model is depicted in Figure 5; the equations are given in Table 3. The state variables are the summer and winter-spring algal phosphorus (A_s, A_w), detritus phosphorus (D), and dissolved reactive phosphorus (P). Measurements showed that zooplankton activity is low in Lake Balaton; hence zooplankton is not modeled explicitly, and the effects of grazing are implicitly incorporated in the algal mortality coefficients (k_{di} , $i = s, w$). Similarly, bacterial effects are incorporated in the mineralization

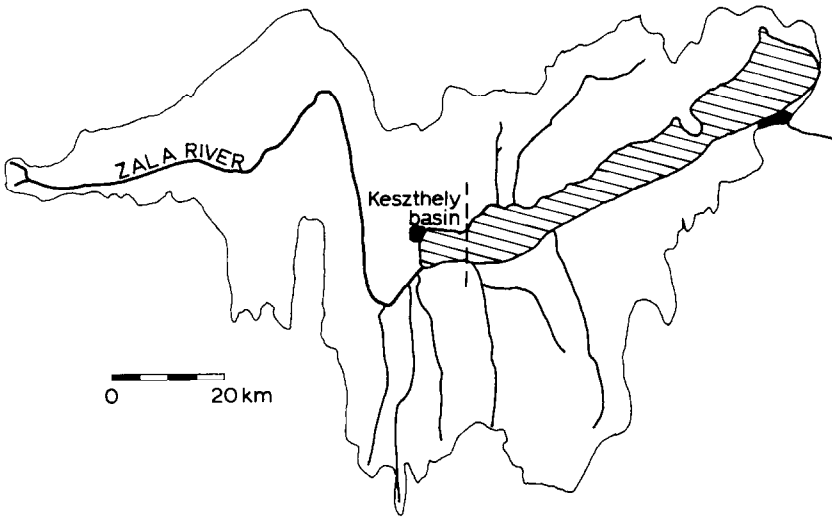


FIG. 2. Lake Balaton, Hungary.

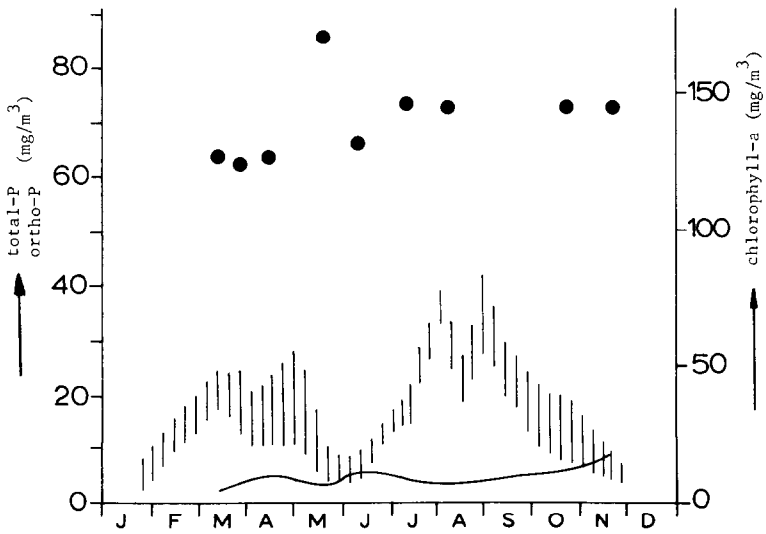


FIG. 3. Observations in the Keszthely Basin, Lake Balaton, in 1977: dots, total P (particulate inorganic P subtracted); dashes, chlorophyll-a; solid line, dissolved reactive P.

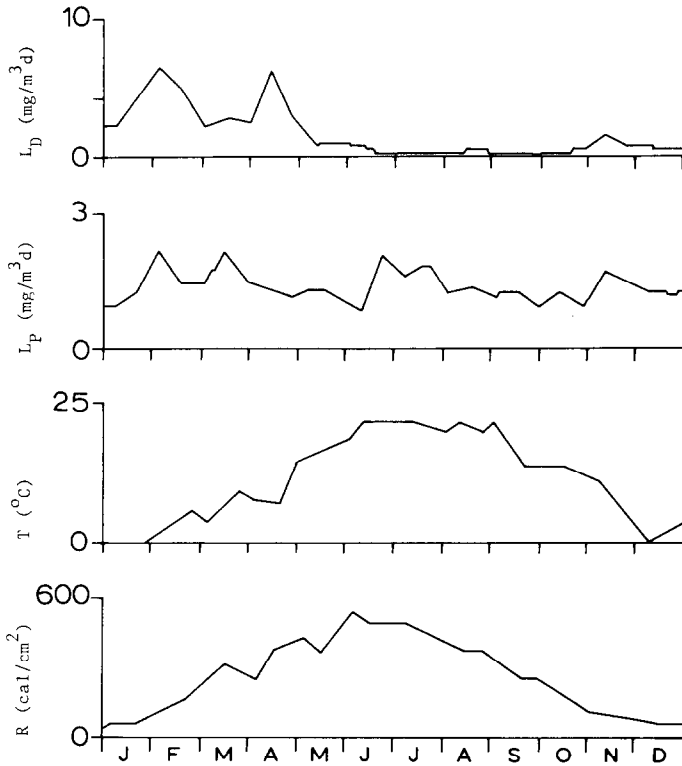


FIG. 4. Forcing functions in 1977.

constant (k_m). The light limitation f_l is the well-known depth and day averaged Steele equation. It is a function of the algal concentration through self-shading. Growth-rate limitation for phosphorus is described by the classical Monod expression f_p . The system is driven by the external loads L_P and L_D , mainly from the Zala River but including contributions of direct sewage, urban runoff, and precipitation. L_D contains dissolved organic P as well as particulate P, and it is assumed that all of it contributes to the mineralizable P fraction, which is called *detritus* in the sequel. Transport of phosphorus occurs through settling of algae and detritus. Since part of the detritus is dissolved (fraction γ), the settling velocity is an apparent settling velocity corrected for the dissolved fraction. This is a primitive model introduced to prevent the need to define the two detritus fractions as separate state variables. In general, there is a net accumulation of P in the lake. Some 10% only is exported with hydrologic throughflow (reciprocal residence time q_v).

TABLE 2
OBSERVATIONAL DATA AND SOME COMPUTED VALUES^a

		Spring		Summer	
		Peak	Average	Peak	Average
Day No.	t	90	60–110	210	165–245
Temperature	T	8	7.1	21	21.5
Irradiation	R	275		430	
Daylength	λ	0.52		0.64	
Chlorophyll-a	y_C	50	40	70	40
Algal P	y_A	25	20	35	20
Light saturation	I_s	165		270	
Light factor	f_l	0.183		0.188	
Detritus load	L_D	2.24 ^b	3.44	0.20 ^b	0.23
Ortho-P load	L_P	1.88 ^b	1.63	1.72 ^b	1.56
Ortho-P	y_P	5	5	5	5
	$\Delta y_P / \Delta t$		0		0
Total P	y_T	70	75	70	70
	$\Delta y_T / \Delta t$		0.3		0
Detritus P	y_D	40	50	30	45
	$\Delta y_D / \Delta t$		-0.3		-0.3

^a Fixed data: $\epsilon_0 = 2.0 \text{ m}^{-1}$, $\alpha = 0.033 \text{ m}^2/(\text{mg P})$, $H = 2.16 \text{ m}$, $r = 0.75$, $I_{s,0} = 68 \text{ cal/cm}^2 \text{ day}$, $c = 3.2 \times 10^{-3} \text{ cm}^2/\text{cal}$.

^b Calculated average over 3 previous weeks.

The accumulation leads to a considerable P pool in the sediment, and the possibility that P is released from the sediments in certain periods of the year is taken into account by L_{int} , the internal source. The options for modeling this term are discussed below.

Parameter Dimensionality Reduction

Despite its relative simplicity from the point of view of biology, the model has no less than 27 parameters and constants (Table 4). The first step is to reduce this number by employing available information. An important source of parameter values in phytoplankton models is primary production measurements. Data collected for Lake Balaton have been processed in a model-oriented analysis [13], leading to estimates for growth rates, temperature dependences, light saturation, and light-saturation adaptation. Slight modifi-

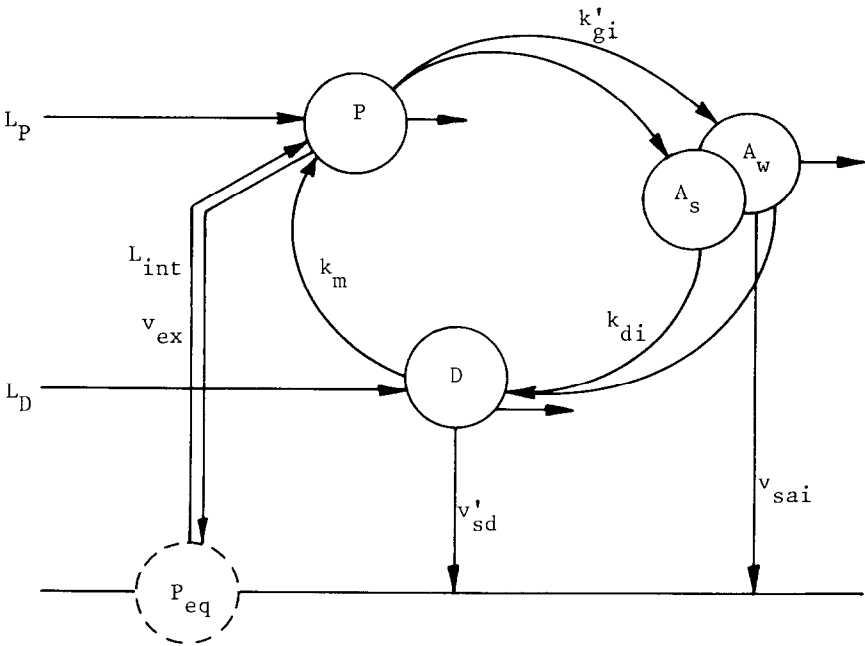


FIG. 5. Fixed cell quatum model structure.

cations were made in the equations, particularly for the rapid adaptation to incident radiation, which turned out to be the same for the summer and winter-spring communities, and hence could be described with the same formula. Table 5 summarizes the parameters included and the values adopted. Direct information was available on water depth and reciprocal residence time. In addition, from published primary-production related measurements [4], an average reflection of 0.75 was estimated, and a relationship between secchi-disk depth and total extinction was derived. By plotting observed secchi-disk depths against chlorophyll-a in 1977, the extinction of the water without algae was estimated as 2.0 m^{-1} , and by assuming a chlorophyll-a/P ratio of 2 an estimate for the self-shading coefficient could be obtained, although the correlation was not very good; most likely the contribution of resuspended sediment particles cannot be assumed to be constant, as was done in the model, and hence the estimated extinction parameters must be considered subject to some stochastic variability.

The Observation Matrix

Since there are four state variables and three observation variables, the observation matrix is not a square matrix, and it is not possible to reconstruct

TABLE 3
FCQ MODEL EQUATIONS

$$\dot{A}_i = k_{gTi} f_{li} f_{Pi} A_i - k_{dT_i} A_i - k_{sai} A_i - q_v A_i, \quad i = s, w$$

$$\dot{D} = \sum_i k_{dT_i} A_i - k_{mT} D - k'_{sd} D + L_D - q_v D$$

$$\dot{P} = - \sum_i k_{gTi} f_{li} f_{Pi} A_i + K_{mT} D + L_{int} + L_P - q_v P$$

$$k_{gTi} = k_{gmi} f_{Ti}, \quad i = s, w$$

Summer algae:

$$f_{Ts} = \begin{cases} \frac{T_{cs} - T}{T_{cs} - T_{os}} \exp\left(1 - \frac{T_{cs} - T}{T_{cs} - T_{os}}\right) & \text{if } T < T_{cs} \\ 0 & \text{if } T \geq T_{cs} \end{cases}$$

Winter Algae:

$$f_{Tw} = \frac{2}{\exp[b(T - T_{ow})] + \exp[-b(T - T_{ow})]}$$

$$k_{dT_i} = k_{di} \theta_{di}^{T-20}, \quad i = s, w$$

$$k_{mT} = k_m \theta_m^{T-20}$$

$$f_{li} = \frac{\epsilon \lambda}{\epsilon H} \left[\frac{1 - \exp(-2R' \exp(-\epsilon H / \lambda I_{si}))}{2R' \exp(-\epsilon H / \lambda I_{si})} - \frac{1 - \exp(-2R' / \lambda I_{si})}{2R' / \lambda I_{si}} \right], \quad i = s, w$$

$$R' = rR$$

$$\epsilon = \epsilon_0 + \alpha \sum A_i$$

$$I_{si} = I_{si0} \exp[c_i R], \quad i = s, w$$

$$f_{Pi} = \frac{P}{P_{ki} + P}, \quad i = s, w$$

$$k_{sai} = v_{sai} / H, \quad i = s, w$$

$$k'_{sd} = v'_{sd} / H$$

$$L_{int} = \frac{v_{ex}}{H} (P_{eq0} \theta_{eq}^{T-20} - P)$$

TABLE 4
TOTAL NUMBER OF PARAMETERS

Rate coefficients	$k_{gms}, k_{gmw}, k_{ds}, k_{dw}, k_m$	5
Settling velocities	$v_{sas}, v_{saw}, v'_{sd}$	3
Temperature parameters	$T_{cs}, T_{os}, b, T_{ow}, \theta_{ds}, \theta_{dw}, \theta_m$	7
Light-function parameters	$I_{ss0}, c_s, I_{stw0}, c_w, \alpha, r$	6
Nutrient-function parameters	P_{ks}, P_{kw}	2
Internal-load parameter	$L_{int}(v_{ex}, P_{eq0}, \theta_{eq})$	1 (3)
External constants	ϵ_0, H, q_v	3
		27 (29)

TABLE 5
PARAMETER REDUCTION

Parameter	Value	Source
k_{gms}	10.8 day ⁻¹	Primary production experiments
k_{gmw}	6.0 day ⁻¹	Primary production experiments
T_{cs}	27.5°C	Primary production experiments
T_{os}	24.5°C	Primary production experiments
b	0.5 °C ⁻¹	Primary production experiments
T_{ow}	8°C	Primary production experiments
$I_{ss0} = I_{sw0} = I_{s0}$	68 cal/cm ² day	Primary production experiments
$c_s = c_{tc} = c$	3.2×10^{-3} cm ² /cal	Primary production experiments
α	0.033 m ³ /(mg P)	Extinction and secchi disk
ϵ_0	2.0 m ⁻¹	Extinction and secchi disk
r	0.75	Extinction and secchi disk
H	2.16 m	Maps
q_c	0.01 day ⁻¹	Water balance

the ratio between summer and winter algae from observations on chlorophyll-a without additional assumptions. Algal counts made in another year [4] suggest that the algal populations in spring and summer are quite different. So for the present rough analysis it is reasonable to assume that only one of the two groups dominates in each period. With the notation

$$y = \begin{bmatrix} y_c \\ y_T \\ y_p \end{bmatrix} \quad \text{and} \quad x = \begin{bmatrix} \Sigma A_i \\ D \\ P \end{bmatrix},$$

where y_c , y_T , y_p represent the chlorophyll-a, the total-P, and the soluble-reactive-P observations, respectively, and ΣA_i is either A_w or A_s , depending on the season, we get

$$H = \begin{bmatrix} \beta & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad (11)$$

with β the ratio of chlorophyll-a to algal phosphorus. Given the observations, the state variables needed for the three types of analysis can be reconstructed using

$$H^{-1} = \begin{bmatrix} 1/\beta & 0 & 0 \\ -1/\beta & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}. \quad (12)$$

From this presentation, an essential weakness of any attempt to model chlorophyll-a from a phosphorus oriented model immediately becomes apparent: the ratio of chlorophyll-a to algal-P must be known, and moreover it must be constant over the period of interest.

Extremum Analysis

A condition for an algal extreme to occur is that $\dot{A} = 0$. From the model equations it follows

$$k_{dT}^* = k_{gT}^* f_I^* f_p^* - k_{sa} - q_v, \tag{13}$$

where the asterisk denotes the value at the time of the extreme. If, from the measurements, $y_A^*, \epsilon_0, \alpha, R^*, \lambda^*, I_s^*$ are known, f_I^* can be computed, and equation (13) specifies a condition for the specific mortality rate, with f_p^* as a parameter.

The mortality rate would even be completely determined if f_p^* could be computed. A problem is that in P-limited lakes P values are near the detection limit, and are therefore very uncertain. So f_p^* would be uncertain even if P_k were known. Table 6 lists the values of interest for the spring and summer peaks, and shows the strong interrelationship between P_k, P^* , and k_{dT}^* .

With respect to the mortality rate, the modeler has two alternatives. He can leave k_{dT} as a calibration parameter. However, if not too much value is to be attached to the P levels computed, the algal peak concentration is not sensitive to k_{dT} . In fact, the model always tends to adapt the P concentration such that equation (13) is fulfilled, and this can be achieved at almost any

TABLE 6
MORTALITY-RATE SENSITIVITY FOR ASSUMED
P LIMITATION LEVEL AT
SPRING AND SUMMER EXTREMES^a

f_p	\tilde{P} at $P_k = 10$	P_k at $\tilde{P} = 5$	k_{dT} (day ⁻¹)	
			$T = 8^\circ\text{C}$ (spring)	$T = 21^\circ\text{C}$ (summer)
0.167	2	25	0.15	0.20
0.231	3	17	0.22	0.28
0.286	4	12.5	0.28	0.36
0.333	5	10	0.33	0.42
0.375	6	8	0.38	0.48

^aParameters used: $k_{gms} = 10.8 \text{ day}^{-1}$ ($k_{gTs} = 7.3$), $k_{gmw} = 6 \text{ day}^{-1}$ ($k_{gTw} = 6$), $v_{sas} = v_{suw} = 0.05 \text{ m/day}$, $v'_{sd} = 0.167 \text{ m/day}$

value of k_{dT} . So an alternative is to fix k_{dT} at 5 to 10% of k_{gT} , and take the calculated P^* for granted.

Period-Average Analysis

A suitable variable for the period-average analysis is total phosphorus. Let P_T denote the total phosphorus according to the model. Then from the model equations it follows that

$$\dot{P}_T = - \sum_i k_{sai} A_i - k'_{sd} D + L_D + L_P + L_{\text{int}} - q_v P_T. \quad (14)$$

Of course, $y_T = P_T$ and the right-hand side of (6) is equal to the average of the right-hand side of (14). Since Equation (14) is linear in x , and the inputs do not occur in the coefficients, Equation (9) applies, leading to the condition

$$k'_{sd} = \frac{\bar{L}_D + \bar{L}_P - q_v \bar{y}_T - \Delta y_T / \Delta t}{\bar{y}_D} - k_{sa} \frac{\bar{y}_A}{\bar{y}_D} + \frac{\bar{L}_{\text{int}}}{\bar{y}_D} \quad (15)$$

where \bar{y}_D is a shorthand notation for

$$\bar{y}_D = - \frac{1}{\beta} \bar{y}_c + \bar{y}_T - \bar{y}_p, \quad (16)$$

and $\bar{y}_A = (1/\beta) \bar{y}_c$. Equation (15) specifies a relationship between the algal settling and apparent detritus settling coefficients and the internal P source, which must be obeyed to yield correct average total P levels in the model.

Equation (15) offers an opportunity to estimate the pair of settling rate coefficients k_{sa}, k'_{sd} (or settling velocities v_{sa}, v'_{sd}) if the additional assumption is made that in the first part of the year the internal source is absent due to the low temperatures. Using the data of Table 2, we obtain (for days 60–110)

$$k'_{sd} = 0.08 - 0.4 k_{sa},$$

or

$$v'_{sd} = 0.17 - 0.4 v_{sa}. \quad (17)$$

It should be noted that the observational-data uncertainty is reflected in this relationship. For example, if y_T is 65 rather than 75, the result is

$$v'_{sd} = 0.22 - 0.5 v_{sa}. \quad (18)$$

This shows the sensitivity of the total P to the value of the detritus settling velocity. In fact, data uncertainties lead to regions in parameter space rather than to strict relationships.

Application of the same analysis for the summer of 1977 (days 165–245) leads to

$$v'_{sd} = 0.052 - 0.44 v_{sa} + 0.048 \bar{L}_{int}. \quad (19)$$

Comparison of Equation (19) with Equation (17) shows that either the apparent settling velocity of detritus has declined, or there has been a considerable internal P source in summer, or both. Without additional information no decision can be made; the analysis hence points towards a focus for additional research. A rough calculation based on measured pore water concentrations of 50–250 mg/m³ yields a release of 0.1–5 (mg P)/m³ day, with a larger probability for the high side of the range in summer [8]. Assuming, for the moment, that the apparent settling velocity of detritus remains the same throughout the year, \bar{L}_{int} must have been on the order of 2.5 mg/m³ day (3.5 mg/m³ day if $v'_{sd} = 0.22$). This is almost independent of the algal settling velocity (possible range 0–0.05 m/day).

The internal P source can be modeled by the simple model

$$\bar{L}_{int} = \frac{v_{ex}}{H} \overline{P_{eq} - P}, \quad (20)$$

where v_{ex} is a transport coefficient, and P_{eq} is determined by chemical adsorption and by the mineralization process in the sediment. Based on adsorption experiments with comparable sediments, Van Raaphorst and Brinkman [10] estimated v_{ex} as approximately 0.05 m/day. With $\bar{P} = 5$ mg/m³ we then have $\bar{P}_{eq} = 105$ mg/m³ (145 mg/m³ respectively), which is consistent with the observations.

Since the internal source is absent in winter and spring, the model equation (20) must be further expanded. One possibility is to make P_{eq} a function of temperature, e.g.

$$P_{eq} = P_{eq0} \theta_{eq}^{T-20}. \quad (21)$$

The physical basis is that mineralization in the sediment is a temperature related process, probably with a strong temperature coefficient due to the acceleration by bacterial biomass increase. By setting $\bar{L}_{int} = 0$ in spring and 2.5 (3.5) mg/m³ day in summer, an estimate of the parameters P_{eq0} and θ_{eq}

can be computed on the basis of actual temperature records. The result is

$$\theta_{eq} = 1.24 \text{ (1.27),}$$

$$P_{eq0} = 75 \text{ (100).}$$

The development above illustrates the usefulness of period-average analysis as a tool for parameter-space delimitation, and also leads to suggestions for model improvement. The uncertainty remaining can only be resolved by further experimental work in the field.

Quasisteady-State Analysis

The steady-state condition for algae has the same form as Equation (13). Two other parameter conditions can be derived by setting the remaining derivatives equal to zero. The results are

$$k'_{sd} = \frac{\bar{L}_D + \bar{L}_p - q_v \bar{y}_T}{\bar{y}_D} - k_{sd} \frac{\bar{y}_A}{\bar{y}_D} + \frac{\bar{L}_{int}}{\bar{y}_D} \tag{22}$$

and

$$k_{mT} = k_{dT} \frac{\bar{y}_A}{\bar{y}_D} - k'_{sd} + \frac{\bar{L}_D}{\bar{y}_D} - q_v. \tag{23}$$

TABLE 7
SPREADSHEET CALCULATION SCHEME OF CALIBRATION PARAMETER ESTIMATES

	Step No.	Inputs	Fixed parameters	Intermediate results	Calibration parameter estimates
Spring	1	$\bar{T}, \bar{y}_c, \bar{y}_T, \bar{y}_p,$ $\delta \bar{y}_T / \Delta t, \bar{L}_D, \bar{L}_p, \bar{L}_{int},$ q_v, H	β, v_{ex}	P_{eqT}	v'_{sd}
	2	t^*, y_A^*, R^*	$\epsilon_0, \alpha, r, I_{s0}$	$\lambda, \epsilon, I_s, f_{Iw}$	
	3	T^*	b, T_{ow}	f_{Tiw}	
	4	y_p^*	P_k, k_{gmw}, θ_d	f_p, k_{gTiw}, k_{dTiw}	k_{dw}
	5	\bar{L}_D, \bar{y}_D		k_{mT}	
Summer	6	$\bar{T}, \bar{y}_c, \bar{y}_T, \bar{y}_p,$ $\Delta \bar{y}_T / \Delta t, \bar{L}_D, \bar{L}_p$		\bar{L}_{int}	θ_{eq}, P_{eq0}
	7	t^*, y_A^*, R^*		$\lambda, \epsilon, I_s, f_{Is}$	
	8	T^*	T_{cs}, T_{os}	f_{Tis}	
	9	y_p^*	k_{gms}	f_p, k_{gTs}, k_{dTis}	k_{ds}
	10	\bar{L}_D, \bar{y}_D		k_{mT}	θ_m, k_m

Equation (22) is very similar to Equation (15), and in fact, application of Equation (15) should be preferred. Equation (23) is of more interest, because it gives a clue to estimation of the mineralization coefficient. It is immediately apparent that assumptions on the mortality rate also effect the mineralization rate. A problem in the application of (23) is the assessment of the steady-state values. The best tactic is to select values observed around the algal peak, or slightly higher, but even so, the estimate obtained cannot be more than a rough first guess. Once the modeler has selected appropriate values for k_d , summer and spring estimates for k_{mT} calculated with Equation (23) can be compared to yield a guess at the mineralization-rate temperature coefficient.

Implementation, Results, and Discussion

Equations (13), (15), and (23) and the associated light and temperature dependences were implemented in a personal-computer spreadsheet program. Both the spring and summer situations were included, and comparison between the two allowed the computation of estimates for the internal load and for the temperature coefficients of the mineralization. The spreadsheet design is shown in Table 7. The parameters in the right-hand column are the results and should be considered as suitable "first guess" values for more formal calibration procedures.

The use of a spreadsheet enables rapid evaluation of how calibration parameters relate to fixed parameters, and makes it easy to judge the effects of observational uncertainty. For example, Figure 6 shows that the various assumptions on summer and spring ortho-P levels, leading to various mortality

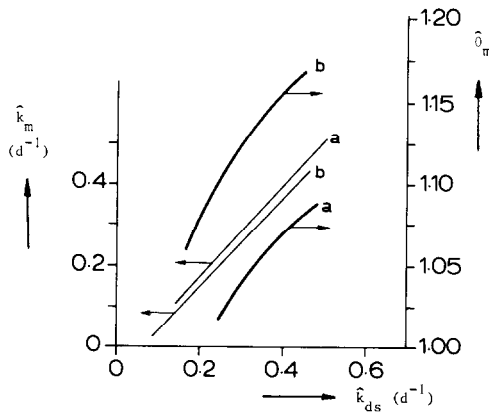


FIG. 6. Interrelationship between summer algal mortality rate ($k_{d,s}$), mineralization rate (k_m), and mineralization temperature coefficient (θ_m): (a) at winter ortho-P = 5 mg/m³, $k_{d,w}$ = 0.76 day⁻¹; (b) at winter ortho-P = 2 mg/m³, $k_{d,w}$ = 0.34 day⁻¹.

rates (Table 6), also lead to the need for adapting the mineralization, in particular the temperature-dependence coefficients. A test by running the model with several combinations from Figure 6 confirmed that they all lead to almost the same simulation result.

Figure 7 shows the simulation with the standard parameter set derived from the spreadsheet. Even without further calibration the concentration levels are in the correct range, but in general the predicted summer algal level is too high. What seems to happen is that the strong temperature dependence of the sediment equilibrium concentration causes an overshoot in internal load during June and part of July when the temperature is higher than average,

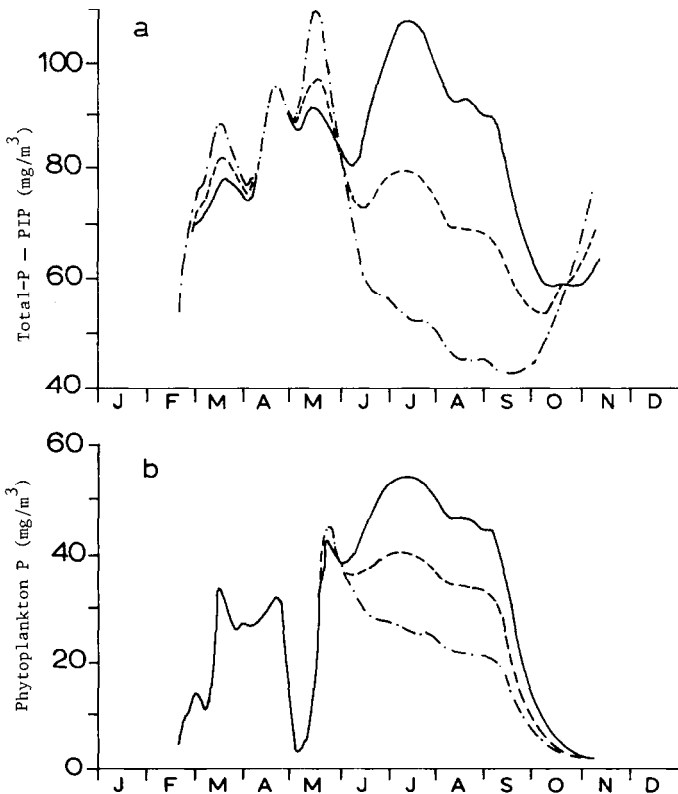


FIG. 7. Phytoplankton P and total P (minus particulate inorganic P): simulation and effect of internal loading. Standard parameter set: $\beta = 2$, $\theta_d = 1.07$, $k_{ds} = 0.40 \text{ day}^{-1}$, $k_{dtw} = 0.76 \text{ day}^{-1}$, $\theta_m = 1.067$, $k_m = 0.39 \text{ day}^{-1}$, $v_{sas} = v_{saw} = 0.05 \text{ m/d}$, $v_{sd} = 0.167 \text{ m/day}$, $P_k = 10 \text{ mg/m}^3$, $\theta_{eq} = 1.25$, $P_{eq0} = 91 \text{ mg/m}^3$, $v_{ex} = 0.05 \text{ m/day}$. Solid line, standard set; dashed line, $v_{ex} = 0.025 \text{ m/day}$, i.e. internal load halved; dot-dash line, $v_{ex} = 0 \text{ m/day}$, i.e. no internal load.

and the only way for the model to compensate is by larger settling and hence larger algal and detritus concentrations. The level cannot be decreased by increasing the mortality rate, because the model adapts its ortho-P level without much change in algal P. Changing the mineralization rate would lead to a shift in the algae/detritus ratio that is not supported by the data. Hence, the only way to influence the algal level is by changing the internal P load.

The role of the internal P source was investigated by running simulations with half the internal load and with no internal load at all ($v_{\text{ex}} = 0.025$ m/day and 0 m/day, respectively). Figure 7 shows the effect. Halving the internal load yields about the correct algal peak levels and average total P. This result indicates that the period-average analysis does not actually guarantee that the correct dynamics are obtained, unless the model is exact. We may also reverse the argument: since the derived period-average balance does not fit, the model must be wrong. That this is so can also be seen from the plots in Figure 7: in contrast with the observations, the model invariably predicts a high algal peak in June, even without internal load. Inspection of the inputs shows that the conditions are favorable for algal growth. The absence of a bloom in June simply cannot be explained by this model.

CONCLUSION

Period-average analysis, extremum analysis, and quasisteady-state analysis are useful aids in reducing the complexity of the parameter-estimation problem in a phytoplankton dynamics model. The analysis confirms once again that curve fitting and model calibration cannot replace process-oriented field research, but the method helps to identify those sectors where further investigations are most needed. The parameter relations obtained reduce the degrees of freedom, and, when supplemented with appropriate uncertainty bounds, they can make formal calibration procedures, including Monte Carlo estimation [3], more efficient. In early stages of model development the need for formal calibration might even be questioned, because analytical techniques will frequently yield satisfactory answers.

LIST OF SYMBOLS

State Variables

\mathbf{x}	State vector
A_i	Algal P concentration; $i = s, w$ for summer, winter, respectively
D	Detritus P concentration
P	Dissolved inorganic P concentration

Observations

y	Observation vector
y_c	Chlorophyll-a concentration
y_A	Algal P concentration
y_D	Detritus P concentration
y_P	Ortho-P concentration
y_T	Total P concentration
H	Measurement matrix, elements h_{ij}

Input Functions or Constants

u	Input vector
R	Day sum of solar irradiation
λ	Length of photoperiod as fraction of day
T	Temperature
L_D	Detritus external load rate
L_P	Inorganic P external load rate
q_v	Flushing rate, i.e. reciprocal residence time
H	Water depth
ϵ_0	Extinction of water without algae

Parameters

p_f	Vector of fixed parameters
p_c	Vector of calibration parameters

Rate Coefficients

k_{gTi}	Specific growth rate at temperature T for species i ; $i = s, w$
k_{dT_i}	Specific mortality rate; $i = s, w$
k_{sai}	Specific settling rate of algae; $i = s, w$
k_{mT}	Specific mineralization rate at temperature T
k'_{sd}	Specific apparent settling rate of detritus

Standard temperature rate coefficients

k_{gmi}	Maximum specific growth rate at T_{os} and T_{ow} for summer and winter algae respectively
k_{di}	Specific mortality rate at 20°C
k_m	Specific mineralization rate at 20°C

Temperature Coefficients

T_{rs}	Critical temperature for summer algae
T_{os}, T_{ow}	Optimal temperature for summer and winter algae, respectively
b	Coefficient for spring algae
θ_{di}	Mortality-rate temperature coefficient
θ_m	Mineralization-rate temperature coefficient

Light-Function Coefficients

ϵ	Total extinction coefficient of water (ϵ_0) and algae
α	Self-shading coefficient
I_{si}	Optimal light intensity for species i ; $i = s, w$
I_{smi}	Basic optimal light intensity; $i = s, w$
I_{sei}	Temperature dependence of optimal light intensity; $i = s, w$

Others

P_{ki}	Monod coefficient for P limited growth; species $i = s, w$
L_{int}	Internal load
P_{eq}	Equilibrium concentration for P exchange
P_{eq0}	Equilibrium concentration at 20°C
θ_{eq}	Temperature coefficient for exchange equilibrium
v_{sai}	Algal settling velocity; species $i = s, w$
v'_{sd}	Apparent detritus settling velocity
v_{ex}	Exchange transport coefficient
β	Chlorophyll-a/(algal P) ratio
*	Value at the extreme
-	Averaged value
-	(Quasi-)steady-state value

REFERENCES

- 1 D. M. Di Toro and G. van Straten, Uncertainty in the parameters and predictions of phytoplankton models, Working Paper WP-79-27, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1979.
- 2 P. Eykhoff, *System Identification—Parameter and State Estimation*, Wiley, Chichester, 1974, 555 pp.
- 3 K. Fedra, G. van Straten, and M. B. Beck, Uncertainty and arbitrariness in ecosystems modelling: A lake modelling example, *Ecol. Modelling* 13(1,2):87–110 (1981)
- 4 S. Herodek and G. Tamas, The primary production of phytoplankton in the Keszthely Basin of Lake Balaton in 1973–1974, *Ann. Biol. Tihany* 42:175–190 (1975).
- 5 R. A. Jones and G. F. Lee, Recent advances in assessing impact of phosphorus loads on eutrophication-related water quality, *Water Resources Res.* 16:503–515 (1982).
- 6 S. E. Jørgensen (Ed.), *Handbook of Environmental Data and Ecological Parameters*, Internat. Soc. Ecol. Modelling, 1979.
- 7 S. E. Jørgensen (Ed.), *Hydrophysical and Ecological Models of Shallow Lakes and Reservoirs*, CP-78-14, IIASA, Laxenburg, 1978.
- 8 L. Lijklema, P. Gelencser, and F. Szilagyi, Sediments and sediment-water interaction, in *Eutrophication of Shallow Lakes: Modeling and Management* (L. Somlyódy, S. Herodek, and J. Fischer, Eds.), IIASA CP-83-53, 1983 pp. 81–100.

- 9 G. T. Orlob (Ed.), *Mathematical Modeling of Water Quality: Streams, Lakes and Reservoirs*, International Series on Applied Systems Analysis No. 12, IIASA-Wiley, 1983, 518 pp.
- 10 W. van Raaphorst and A. G. Brinkman, The calculation of transport coefficients of phosphate and calcium fluxes across the sediment-water interface, from experiments with undisturbed sediment cores, presented at IAWPRC Biannual Symposium, Amsterdam, Sept. 1984.
- 11 L. Somlyódy, Modelling a complex environmental system: The Lake Balaton Study; *Math. Modelling* 3:481–502 (1982).
- 12 L. Somlyódy and G. van Straten, *Modeling and Managing Shallow Lake Eutrophication (with Application to Lake Balaton)*, Springer, 1985.
- 13 G. van Straten and S. Herodek, Estimation of algal growth parameters from vertical primary production profiles, *Ecol. Modelling* 15: 287–311 (1982).
- 14 G. van Straten, Analysis of model and parameter uncertainty in simple phytoplankton models for Lake Balaton, in *Progress in Ecological Engineering and Management by Mathematical Modelling* (D.M. Dubois, Ed.), Editions CEBE-DOC, Liege, Belgium, 1981, pp. 107–134.
- 15 G. van Straten, Lake eutrophication models, in *Eutrophication of Shallow Lakes: Modelling and Management* (L. Somlyódy, Ed.), IIASA CP-83-S3, Internat. Inst. for Appl. Systems Anal., Laxenburg, Austria, 1983, pp. 175–206.
- 16 R. V. Thomann, D. M. Di Toro, R. P. Winfield, and D. J. O'Connor, *Mathematical Modeling in Lake Ontario, 1: Model Development and Verification*, Environmental Protection Agency Ecological Research Series, EPA-660/3-75-005, 1975, 177 pp.
- 17 R. A. Vollenweider and J. J. Kerekes, The loading concept as basis for controlling eutrophication. Philosophy and preliminary results of the OECD programme on eutrophication, *Progress in Water Technology (Norway)* 12:5–38 (1980).