UNCERTAINTY AND ARBITRARINESS IN ECOSYSTEMS MODELLING: A LAKE MODELLING EXAMPLE

K. Fedra, G. van Straten, and M.B. Beck
International Institute for Applied Systems Analysis, Austria

Research Reports, which record research conducted at IIASA, are independently reviewed before publication. However, the views and opinions they express are not necessarily those of the Institute or the National Member Organizations that support it.

FOREWORD

In recent years there has been considerable interest in developing models for river and lake ecosystems, much of it directed toward large and complex simulation models. However, this trend has given rise to a number of concerns, notably those of accounting for the effects of uncertainty and of establishing model validity and credibility.

The International Institute for Applied Systems Analysis is addressing such concerns in its work on environmental quality control and management, one of the principal themes being to develop a framework for modeling poorly defined environmental systems.

This paper discusses, for situations in which the available field data are sparse and uncertain, the problem of constructing, evaluating, and applying a model for prediction -- and ultimately management -- purposes (see also K. Fedra, "Mathematical modelling -- a management tool for aquatic ecosystems?" *Helgoländer Meeresuntersuchungen* 34:221--235, 1980, also reproduced as IIASA RR-81-2). In particular, it emphasizes the close relation between the process of model calibration and the predictions obtained subsequently (see also M.B. Beck, "Hard or soft environmental systems?" *Ecological Modelling* 11: 233--251, 1981, also reproduced as IIASA RR-81-4).

Thus, uncertainty and the reliability of models and forecasts are key concerns of this paper.

JANUSZ KINDLER
Chairman
Resources and Environment Area

UNCERTAINTY AND ARBITRARINESS IN ECOSYSTEMS MODELLING: A LAKE MODELLING EXAMPLE

K. FEDRA, G. VAN STRATEN and M.B. BECK

International Institute for Applied Systems Analysis, A-2361, Laxenburg (Austria)

(Accepted for publication 13 January 1981)

ABSTRACT

Fedra, K., van Straten, G. and Beck, M.B., 1981. Uncertainty and arbitrariness in ecosystems modelling: a lake modelling example. *Ecol. Modelling*, 13: 87–110.

Mathematical models of ecosystems are considerable simplifications of reality, and the data upon which they are based are usually scarce and uncertain. Calibration of large complex models depends upon arbitrary assumptions and choices, and frequently calibration procedures do not deal adequately with the uncertainty in the data describing the system under study. Since much of the uncertainty and arbitrariness in ecological modelling is inevitable, because of both practical as well as theoretical limitations, model-based predictions should at least reveal their dependence on, and sensitivity to, uncertainty and arbitrary assumptions.

This paper proposes a method that explicitly takes into account the uncertainty associated with data for modelling. By reference to a partly qualitative and somewhat vague definition of system behaviour in terms of allowable ranges, an ensemble of acceptable parameter vectors for the model may be identified. This contrasts directly with a more conventional approach to model calibration, in which a quantitative (squared-error) criterion is minimized and through which a supposedly 'unique' and 'best' set of parameters can be derived. The ensemble of parameter vectors is then used for the simulation of a multitude of future systems behaviour patterns, so that the uncertainty in the initial data and assumptions is preserved, and thus the predicted future systems response can be interpreted in a probabilistic manner.

INTRODUCTION

Mathematical models of ecosystems are inevitably considerable simplifications of reality. The data available for analysis are usually scarce, they have mostly been measured infrequently, and are subject to high levels of sampling error. Furthermore, the stochastic variability of the systems themselves is also reflected in the data. Consequently, such data allow only an uncertain description of systems behaviour to be developed; the data are probably

inadequate for the calibration of large complex models. At the same time, more or less arbitrary choices have to be made about the structure of the model, the criteria to be used for model calibration and parameter estimation, and the calibration procedure itself. Such choices are neither unique nor absolute in some sense, yet they will affect the results obtained from the application of the calibrated model. Since it is necessary to make arbitrary assumptions and to use uncertain field data, the authors argue that our modelling results, i.e. model-based predictions about the future, should reveal their dependence on and sensitivity to such assumptions and uncertainty.

Recently, formalized parameter calibration routines have begun to be applied in the field of modelling complex aquatic ecosystems, e.g. by Lewis and Nir (1978), Jørgensen et al. (1978), Di Toro and van Straten (1979), and Benson (1979). In these methods a loss function is defined, usually in a squared-error form, and, subsequently, a parameter vector is sought that minimizes this loss function. This procedure thus avoids the analyst's subjective perception of which parameter ought to be adjusted to improve the fit inherent in the more commonly applied trial-and-error calibration procedure. Also, the equally subjective judgement of agreement between simulation and observation is replaced by a more formal quantitative notion. However, although frequently called 'objective function', this does not imply that the criterion chosen is free from subjective elements. For example, in problems with state variables having different physical dimensions some (subjective) form of weighting is required in the formulation of a single-valued loss function. Furthermore, it is not easy to account for uncertainty in the field data, although methods to do this have been attempted (Beck and Young, 1976; Jolankai and Szöllősi-Nagy, 1978; Lewis and Nir, 1978; Beck, 1979; Di Toro and van Straten, 1979; Fedra, 1979, 1980, 1981a). Finally, however, it has to be recognized, that the assumption that a single 'best' parameter vector exists is at least questionable, especially if data uncertainty is considered, and in any case experience shows that it is extremely difficult to find such a unique vector if the number of parameters to be estimated is larger than, say, six to ten.

This paper proposes a method that explicitly takes into account the uncertainty associated with data for modelling, including initial conditions and forcing functions; the method proposed also circumvents the problem of assuming parameter uniqueness. Basically, instead of assuming the existence of a 'best, unique' parameter vector, which may be found through minimization of a loss function, the method allows a set of vectors to be identified by reference to a more 'vague' definition of systems behaviour. The uncertainty in the available information is expressed in this 'vague' definition, through the specification of bounds between which 'acceptable' model simulations

should fall. Clearly, given these ranges, more than one vector exists that fulfills the requirements of being 'acceptable'. This part of the work owes much to the recently reported study of Spear and Hornberger (1980), although the emphasis and focus are different.

Once it is recognized that the uncertainty of the field data suggests the specification of bounds on acceptable system behaviour, and nothing more precise, it ought also to be recognized that those parameter vectors that are found to give the defined behaviour are all equally good in view of the data uncertainty. Any sample parameter vector that is found to simulate the defined system behaviour can, since it has passed the 'test' of 'calibration', be applied in principle to generate future systems responses under changed conditions (represented by a change in any of the vector elements). Hence a multitude of simulated future system behaviour patterns can be generated from the set of acceptable parameter vectors.

Two points are worth noting about this approach to model-based predictions. First, the close interdependence of estimation and prediction is emphasized (see also Beck, 1980, Fedra, 1981b). Second, the uncertainty in the data has not been ignored in applying the model to the problem of prediction. Uncertainty about the system's behaviour is included explicitly and in effect 'preserved'; thus future responses of the system are predicted in the form of sampled probability distributions, as opposed to unique, average, and deterministic forecasts (see also O'Neill and Gardner, 1978; Tiwari et al., 1978; Di Toro and van Straten, 1979; Beck et al., 1979; Halfon and Beck, 1980; Fedra, 1979, 1980, 1981a).

We shall illustrate the method with an application to a specific problem relating to an Austrian lake (Attersee, Salzkammergut).

METHOD

The method used is estimation and prediction by Monte Carlo simulation. Let us suppose that a given structure for the model is assumed, i.e. we admit, an unavoidable (arbitrary) assumption that will subsequently affect the predictions obtained. Further, let us represent this model by a vector function $\mathbf{f}$ with domain $\mathcal{D}(f)$ and range $\mathcal{R}(f)$. If RD is a subset of $\mathcal{R}$, then the inverse image of RD under $\mathbf{f}$ is the subset of $\mathcal{D}(\mathbf{f})$ given by

$$\mathbf{f}^{-1}(RD) = \{x : \mathbf{f}(x) \in RD\}$$

which we will call CM (see Fig. 1). To identify CM, which in our terminology is the set of all character vectors (for a definition, see below) resulting in the defined realistic model response (RD), we proceed as follows.

For the first step in the approach the system behaviour is defined in terms of the model to be used by a series of constraint conditions. From the field data typical features of the system's behaviour are derived. We consider it
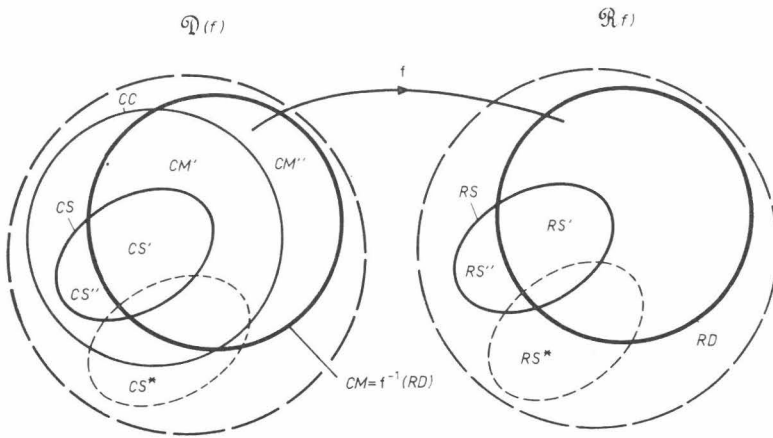
Fig. 1. Set diagram of the relationships of character- and response-space. $\mathcal{D}(\mathbf{f})$: set of all possible character vectors (domain of $\mathbf{f}$); $\mathcal{R}(\mathbf{f})$: set of all possible model responses (range of $\mathbf{f}$); $\mathbf{f}$: model (vector function); CD: defined set of plausible character vectors; RD: defined realistic response region; CM: inverse image of RD; CS: character vectors sampled in Monte Carlo procedure; RS: direct image of CS; CS': subset of CS, generating plausible realistic response RS'; CS'': subset of CS resulting in unrealistic response RS''; CS^*: modified CS' used for prediction, resulting in RS^*.

significant that the behaviour definition, i.e. the constraint conditions, includes both dynamic and aggregate features. The latter, e.g. yearly primary production, are usually more reliable than individual data at any single instant in time, thus they allow for a definition of the system's behaviour in which one intuitively would place more confidence. The set of m ranges of behaviour-describing measures, defines an m-dimensional box in the range-space of the model, or a set RD of plausible, acceptable model responses.

Secondly, from empirical evidence and from previously quoted values for model coefficients, it is possible to specify ranges of inputs, forcings, initial conditions and model coefficients required by the particular model structure. A vector of numerical values taken from these ranges fully characterizes the response of the model. In order to make a distinction from the usual term-parameter vector, we shall call any sample combination from these ranges a character vector. The ranges of n character vector elements define a region CD (the set of all allowable character vectors) in the n-dimensional domain space of the model.

Thirdly, this character vector region CD is now randomly sampled N times by a Monte Carlo technique. Each sample character vector CS_i $(i = 1, \ldots, N)$ is then used for a simulation run, and the resulting set RS of model responses RS_i $(i = 1, \ldots, N)$ is classified according to the behaviour definition RD

behaviour:

$$RS' = \{RS_i : (RS_i \in RD)\} \qquad n(RS') = M$$

non-behaviour:

$$RS'' = \{RS_i : (RS_i \notin RD)\} \qquad n(RS'') = N - M$$

The set CS of sample character vectors is separated correspondingly into two complementary subsets CS' and CS'', with M and $N - M$ elements, respectively. In other words, by this 'calibration' procedure we are looking for a separation of the sample of character vectors into a behaviour-giving subset (CS') and a non-behaviour-giving subset (CS''). The total sample of M plausible behaviour-generating character vectors is then analysed to give some insight into possible relations and interdependencies or the character space configuration.

The character vectors CS_i' giving rise to a response RS_i' completely within the behaviour-defining boundaries are considered as random samples from a character space region CM corresponding to the defined behaviour region RD of the model

$$CS' \subset CM$$
$$CM = f^{-1}(RD)$$

as illustrated in Fig. 1. It should be noted that CM may not be fully included in CD so that there are character vectors, some of whose elements are outside the specified ranges, that give rise to the defined behaviour.

Finally, the set CS' of M sample character vectors CS_i', being identified as 'acceptable' character vectors, is then used for computations of model responses under changed conditions. That is to say, one or more elements of the M character vectors are changed according to the extent of the assumed alteration, and the set of M modified vectors is used to generate probability distributions of model responses. This procedure can be repeated to represent different conditions, and in fact the modification of the vector elements may be done systematically to investigate a range of possible future changes and their significance in terms of response probability distributions.

APPLICATION

Initialization

The lake system
In cooperation with the Austrian Lake Eutrophication Program, Project Salzkammergutseen, the Attersee, a deep, stratified, oligotrophic lake of almost 4000×10^6 m^3 and a theoretical retention time of seven to eight years,

TABLE I

Attersee: basic lake data (after Flögl, 1974)

Geographic position	47°52′N, 13°32′E	Length	20 km
Catchment area	463.5 km^2	Average width	3 km
Surface area	45.9 km^2	Total shore length	53 km
Maximum depth	171 m	Retention time	7–8 y
Mean depth	84 m	Average outflow	17.5 m^3 s^{-1}
Volume	3934×10^6 m^3		

was subjected to our approach. Basic lake data are compiled in Table I. The investigations on the lake, carried out since 1974, and initially within the frame of the OECD Lake Eutrophication Program, indicated an increasing eutrophication of the lake. Increasing phytoplankton peak biomass and decreasing transparency of the water signalled a trend towards eutrophication; however, the variability in the measurements and the comparatively short time-span of observations make it difficult to identify significant changes. Nevertheless, a preliminary study of the nutrient-loading/ production relations seemed to be promising (here primary production per unit lake area was taken as an approximate measure for the trophic state of the lake).

Although the problem setting is somewhat diffuse from the point of view of possible management and water quality control measures, two principal features of the lake system allow us to address practical problems. First, the major nutrient input stems from one point source, namely the upstream mesotrophic Mondsee (14.2 km^2 surface area, 510×10^6 m^3, 247 km^2 catchment area, and a retention time of about two years). More than 50% of the phosphorus loading of the Attersee is attributable to the Mondsee discharge, and a very high fraction of particulate phosphorus is contained in this discharge (Müller, 1979). The impact of possible changes in the trophic state of the Mondsee on the Attersee is, therefore, of considerable interest. Second, a sewer system and associated treatment plant for the sewage discharge to the Attersee have been constructed recently (Flögl, 1974). Again, the impact of these installations on the water quality of the Attersee are of interest.

The data available for our analysis, comprising estimates of nutrient inputs and outputs as well as lake nutrient concentrations, were found to show a high degree of variability both within and between years, and this variability was especially pronounced for the phosphorus measurements. This is, at least in the case of the orthophosphate, due to its low concentrations around 1 mg m^{-3}, which is approximately at the same level as the absolute measurement error. This also led us to the simplifying assumption of a horizontally completely mixed water body; the data would not support a

more detailed spatial resolution for the model. It was also decided to add the observations of several years together to obtain a picture of a typical Attersee behaviour pattern.

The simulation model

For the simulation of the nutrient-loading/production relationship on a yearly time-scale the dynamic lake phosphorus model by Imboden and Gächter (1978) was chosen. The model predicts primary production per unit lake area as related to imports of soluble reactive, as well as particulate phosphorus, various forcings, and model parameters. The relationship between loading and primary production is described by means of a dynamic, one-dimensional, vertical (multi-layer) diffusion model for the two state variables of particulate phosphorus and soluble reactive phosphorus concentrations. The model uses Michaelis–Menten kinetics and self-shading by algae, together with a production rate that varies in time according to the seasonal variations in irradiance and water temperature. Respiration, sedimentation, stratification with vertical eddy diffusivity and variable thermocline depth, lake morphometry, and finally, hydraulic loading are all accounted for in the model. A homogeneous, well-mixed epilimnion is assumed, and phosphorus export is determined by its epilimnion concentration and by hydraulic loading. Zooplankton is not explicitly included in the model; its effects on phytoplankton are included in the first-order loss term describing respiration/remineralization. Consequently, the model is designed more for the simulation of yearly aggregate features than for the simulation of short-term algal population dynamics.

Some minor modifications of the model were made in order to allow for a parameterized description of time-varying forcing functions (production rate and thermocline depth). Rather than specifying these coefficients in the form of tables, as was done originally for the model, we approached the time patterns by simple analytical functions of time. Thus, the dynamic pattern of the production rate is described by a sine function with the minimum, maximum, and the time of the maximum as auxiliary parameters. Similarly, the thermocline depth is a linear function of time, characterized by the depth and time at the outset of stratification and the depth and time at the end of the stratified period. Other potentially time-varying data (e.g. nutrient inputs, hydraulic loading, eddy coefficients) were kept constant, since the available field data did not allow a meaningful yearly pattern to be specified. In view of the morphology of the Attersee, the backflux of phosphorus from the sediments was set to zero in the model.

Ultimately, a total of 22 character vector elements (parameters, forcing function-related parameters and initial conditions) were required in this application. These are listed by name in Table II, together with the ranges

TABLE II

Character vector elements and the ranges used

Data type	Unit	Minimum	Maximum
Parameters *sensu stricto*			
1. Michaelis constant (phosphorus)	$mg\ m^3$	0.20	2.00
2. Respiration/mineralization epilimnion	d^{-1}	0.02	0.20
3. Respiration/mineralization hypolimnion	d^{-1}	0.01	0.025
4. Net sedimentation velocity epilimnion	$m\ d^{-1}$	0.01	0.75
5. Net sedimentation velocity hypolimnion	$m\ d^{-1}$	0.025	2.00
6. Diffusion coefficient hypolimnion	$cm^2\ s^{-1}$	0.02	0.50
7. Diffusion coefficient thermocline	$cm^2\ s^{-1}$	0.01	0.25
8. Extinction coefficient	m^{-1}	0.20	0.40
9. Self-shading coefficient	$m^2\ mg^{-1}$	0.01	0.02
10. Thickness of thermocline	m	5.00	10.00
Import and forcing describing data			
11. Orthophosphate import	$mg\ m^{-2}\ d^{-1}$	0.01	0.20
12. Particulate phosphorus import	$mg\ m^{-2}\ d^{-1}$	0.25	1.50
13. Hydraulic loading	$m\ d^{-1}$	0.03	0.05
14. Minimum production rate	d^{-1}	0.25	0.50
15. Maximum production rate	d^{-1}	1.00	10.00
16. Time lag of production maximum	d	180	270
17. Initial thermocline depth	m	3.00	6.00
18. Final thermocline depth	m	15.00	20.00
19. Beginning of stratified period	day	120	280
20. End of stratified period	day	280	330
Initial conditions			
21. Initial orthophosphate mixed period	$mg\ m^{-3}$	0.20	2.00
22. Initial particulate P mixed period	$mg\ m^{-3}$	2.50	7.00

used in the Monte Carlo simulation. The minimum and maximum values, which define the ranges, were obtained either from the known variability of available estimates (e.g. particulate phosphorus loading) or from expansion around values given in the literature. It should be pointed out that the results of the method are not critically influenced by the ranges selected, as long as they are ecologically or physically feasible. However, reduction of the ranges wherever possible is useful for increasing the efficiency of the computation. Thus, for several of the character vector elements the ranges given in Table II were obtained after reduction on the basis of an initial set of 10 000 pilot runs.

The simulation model was incorporated as a subroutine in a control program, that generated random sample character vectors from the ranges specified. Since a priori information on the probability distributions and correlation structure of the parameters was absent, independent rectangular

distributions were assumed. For each character vector, one simulation run was completed (for a period of one year) and the model response stored for subsequent analysis.

Behaviour definition

The output of any given model run has to be compared with the (defined) system behaviour in order to enable classification of the character vectors into a behaviour-giving class and a non-behaviour-giving class. It is obvious that the definition of the system's behaviour is a crucial step in the analysis. The system behaviour definition should reflect all the available, relevant knowledge on the system (e.g. that obtained from field observations). It is worth noting that a definition of system behaviour derived from the observations does not depend upon the model. However, this definition has to be specified in terms of model output constraint conditions, or, in other words, the behaviour definition must be cast within the framework of the model actually used.

Based on the available data and discussions with biologists acquainted with the Attersee system, seven indices were selected for inclusion in the behaviour definition. The constraints placed on these indices for the purpose of behaviour definition were specified such that the measurement uncertainty and the natural stochastic variability of the ecosystem (including variability among the years) were accounted for. Table III summarizes the resulting behaviour definition.

In this way, the behaviour definition can be viewed as a seven-dimensional box in the response space and a model simulation run has to lie completely within this box in order to be classified as a simulation belonging to the behaviour-giving class.

TABLE III

Definition of the system's behaviour

---

1. Total primary production per year has to be between 50 and 150 g C m^{-2}.
2. Total phosphorus export per year has to be between 2 and 8 tons.
3. The peak value of particulate phosphorus in the epilimnion has to occur between Julian day 60 and Julian day 210.
4. The peak value of particulate phosphorus in the epilimnion must not exceed 15 mg P m^{-3}.
5. The concentration of orthophosphate during the mixed period must not exceed 2.5 mg P m^{-3}.
6. The peak value of particulate phosphorus must be at least twice the minimum value.
7. The maximum total phosphorus content of the lake during the year must not exceed twice the minimum value.

Analysis

Character vector space structure

Out of 10000 runs only 293 character vectors were found that gave rise to a model output fully within the behaviour constraint conditions given in the previous section. Inspection of the sample ranges of individual elements of these 293 vectors showed that no further 'rectangular' reduction of the character vector space (CD) was possible. In other words, the boundaries of the behaviour-giving character vector space region (CM), as sampled by the 293 behaviour-giving vectors (CS') extended up to the boundaries of the 22-dimensional character vector box (CD). Figure 2 shows the distribution of the behaviour-giving class for one of the character vector elements in order to illustrate this point. Figure 2 also suggests that there are regions in the character vector space (CD) where one is more likely to find a behaviour response than in others.

Due to the high dimensionality of the character vector space the geometry of the behaviour-giving region is in general difficult to investigate. However, a tentative exploration of the distributions of Fig. 2 can be made by projecting these distributions onto a two-dimensional surface. An example is given in Fig. 3. From this figure it is apparent that the behaviour-giving parameter combinations are more densely clustered in certain regions. It is also evident that other regions in the 22-dimensional box are 'empty'.
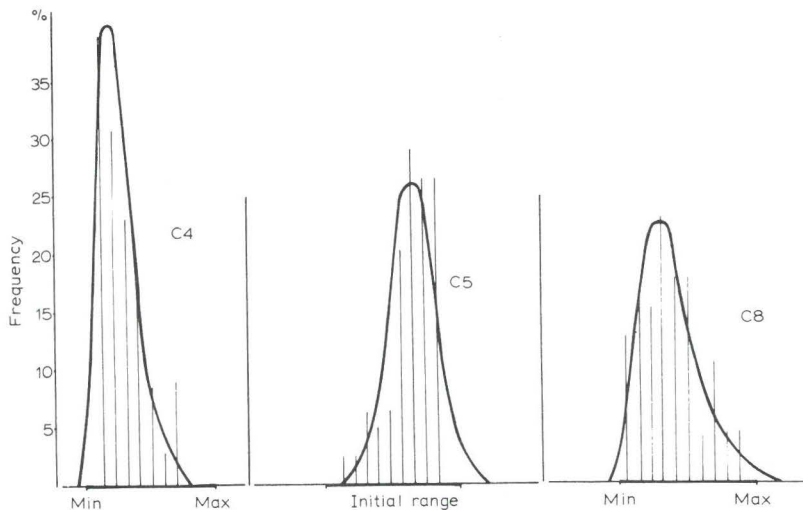

Fig. 2. Frequency distributions of behaviour-giving values of individual character vector elements within the ranges sampled (compare Table II), and probability distributions fitted; C4: sedimentation rate epilimnion; C5: sedimentation rate hypolimnion; C8: extinction coefficient.
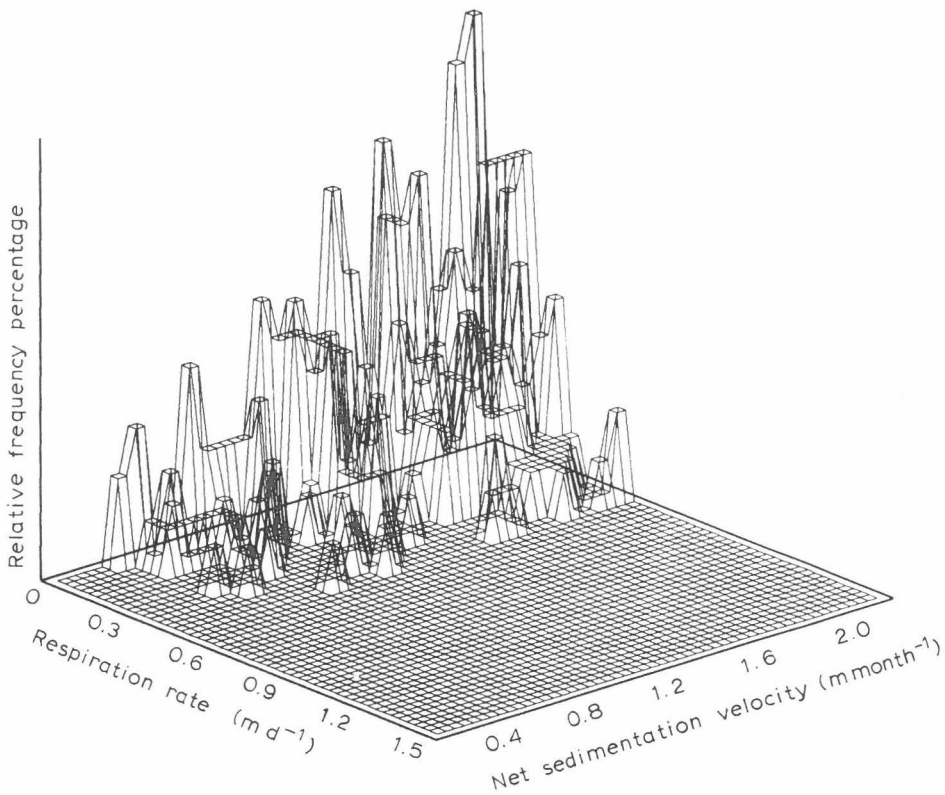
Fig. 3. Projection of character vector space for the behaviour-giving ensemble; combinations of two character vector elements over their initial ranges used for sampling.

However, in view of the large number of non-behaviour-giving combinations, we must conclude that almost every individual value of a character vector element can give rise to the behaviour or not, depending on the sample values of the other elements. Thus, as also suggested by Fig. 3, it is the (multiple) correlations between the character vector elements that determine the shape of the behaviour-giving character space.

To gain insight into the structure of the model (and, hopefully, of the system) a correlation analysis was performed. Consequently, 13 out of the 22 character elements were found to be significantly correlated with one or more of the other character elements. The most complex relations, with four or five significant pairwise correlations, were found for the respiration/mineralization rate in the hypolimnion, the net sedimentation velocities, the particulate phosphorus import, and the hydraulic loading. Also, production rate maximum and time lag showed more than one significant correlation. Correlations between the character vector elements of the behaviour-giving class reflect the ability of the model to balance one extreme with another,
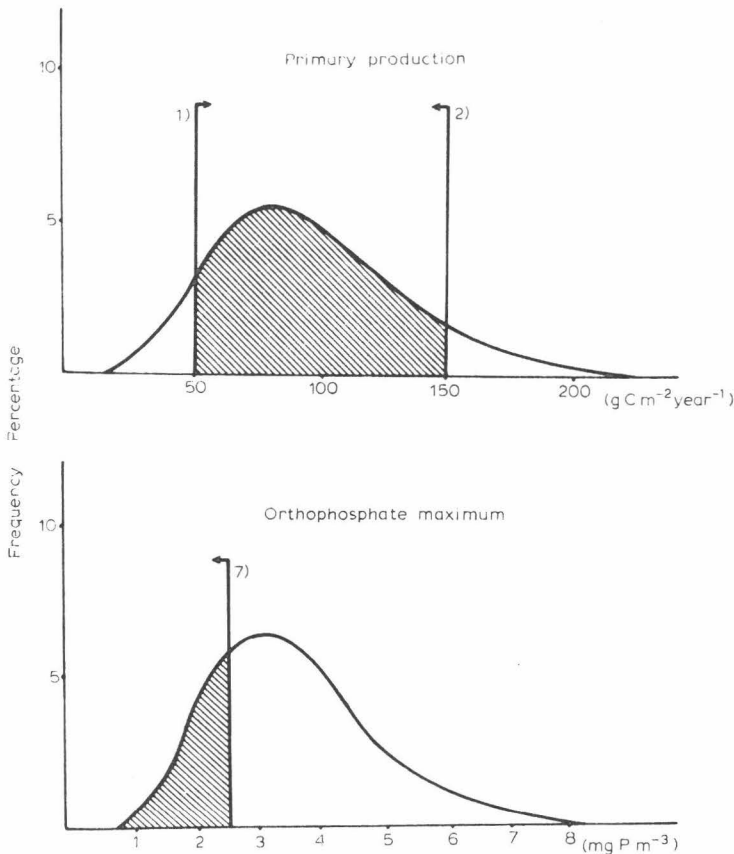
Fig. 4. Frequency distributions of individual model response variables used for the definition of realistic behaviour; constraint conditions delimiting the realistic range are indicated. Model response to 10000 independent random samples from the character space *CD*.

while still fulfilling the behaviour definition constraints. Character vector elements which would force the behaviour-defining variables in the same 'direction' (relative to the boundaries of the seven-dimensional behaviour box) can be expected to be negatively correlated, and vice versa for the positive correlations. In this way, e.g. the strong positive correlation of particulate phosphorus import and sedimentation velocity (epilimnion), it indicates that the constraint variables yearly primary production and/or the allowable algal biomass peak value are sensitive to the 'net' effect of these counteracting processes. The constraint of maximum allowable orthophosphate concentration provides another example. Hypolimnic remineralization, a major process affecting the value of orthophosphate concentration, is consequently negatively correlated with orthophosphate import and with the initial phosphorus concentrations.

In geometrical terms it can be said that the correlations indicate the main axes along which the behaviour-giving character vector region (*CM*) is oriented. Consequently, the model response (of giving the defined behaviour) is most strongly influenced by varying the parameter values in a direction orthogonal to these axes. In this sense the correlation matrix can be interpreted in terms of a sensitivity analysis as well.

According to the character vector correlation structure, the Attersee system, as defined by its geomorphology and the behaviour definition, is characterized by a delicate balance between the processes responsible for primary production and phosphorus export (which is mainly determined by the epilimnic phosphorus concentration) and those which determine the orthophosphate peak concentrations, namely (besides the imports), sedimentation to the comparatively large hypolimnion and hypolimnic remineralization. This balance can only be achieved with a high phosphorus turnover in the epilimnion and comparatively slow net remineralization in the hypolimnion. For a lake in the geographical position of the Attersee and with Attersee's morphometric features and associated temperature distributions, this seems to be a reasonable interpretation.

The role of the behaviour definition

Choosing values for the constraints on the allowable behaviour patterns is subject to a two-sided problem. On the one hand, the constraint ranges should be sufficiently narrow so as to restrict the allowable patterns such that they represent unambiguously the system's empirical behaviour in a meaningful way. On the other hand, all the variability in the system behaviour and the uncertainty in the observations should be taken care of with a minimum of arbitrariness. Since a reconciliation of these two objectives is rather difficult in practice, we examined the effects of the constraint setting on the character vector classification. For this purpose, the model response space was projected onto the individual model output variable axes. The position of the constraints in relation to the resulting frequency distributions (Fig. 4) gives some indication of the relative importance of the individual constraints. Figure 5 shows two contrasting examples for a pair of critical and uncritical constraints, respectively.

The original constraint values were altered and the effects of this on the character vector separation were studied by logging violations of the constraints. For the standard set of constraint values (Table III) the numbers of violations together with a relative coincidence matrix of violations are shown in Table IV. Clearly, the allowable phosphate level and the first permissible day for the algal peak are the major constraints on achieving an overall 'acceptable' model response. Some of the other constraints are either not violated at all, e.g. minimum relative biomass increase, or are rarely violated,
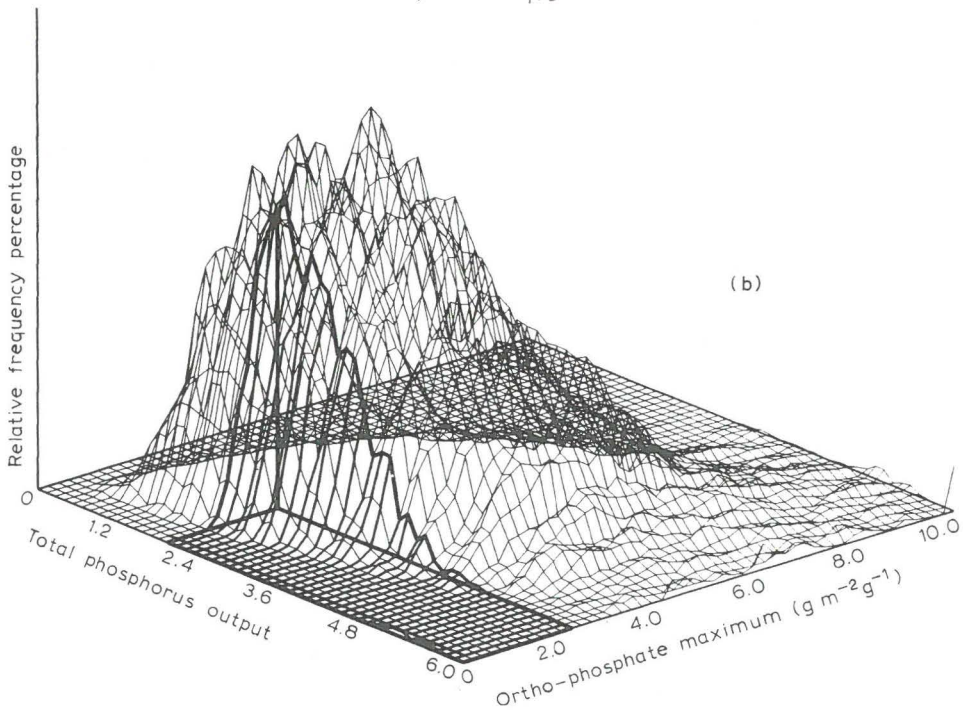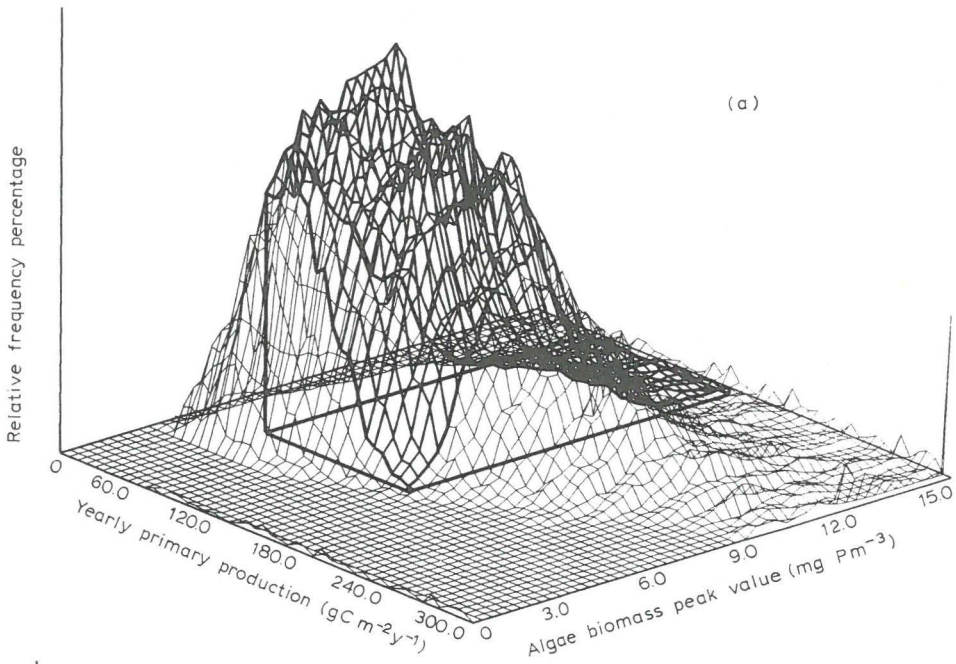
Fig. 5a. Model response space projections on planes of two response variables, indicating the position of the constraint conditions for a pair of uncritical conditions.

Fig. 5b. Model response space projection on planes of two response variables, indicating the position of the constraint conditions for a pair of critical conditions.

TABLE IV

Constraint violations

(Standard definition; 10 000 runs)
1. Primary production too low   1237 cases
2. Primary production too high   904 cases
3. Biomass peak too early   5108 cases
4. Biomass peak too late   1492 cases
5. Biomass peak too high   357 cases
6. Relative biomass increase too low   0 cases
7. Phosphate level too high   7201 cases
8. Phosphorus export too low   2398 cases
9. Phosphorus export too high   1 case
10. Relative change in P-content too high   2148 cases

Percent coincidence matrix

Condition	1	2	3	4	5	6	7	8	9	10
1	100.0	0.0	77.8	8.2	28.5	0.0	35.6	67.7	0.0	17.5
2	0.0	100.0	13.3	39.5	0.4	0.0	100.0	0.0	0.1	9.5
3	18.8	2.4	100.0	0.0	6.9	0.0	60.5	36.5	0.0	34.1
4	6.8	23.9	0.0	100.0	0.3	0.0	89.1	8.1	0.1	13.7
5	98.9	1.1	98.9	1.1	100.0	0.0	1.1	98.9	0.0	0.6
6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	6.1	12.6	42.9	18.5	0.1	0.0	100.0	10.1	0.0	13.6
8	34.9	0.0	77.7	5.1	14.7	0.0	30.2	100.0	0.0	36.0
9	0.0	100.0	0.0	100.0	0.0	0.0	100.0	0.0	100.0	0.0
10	10.1	4.0	81.0	9.5	0.1	0.0	45.7	40.2	0.0	100.0

e.g. the upper limit of total phosphorus output. There are also some notable relationships in the violations observed. For example, in almost all cases in which condition 5 is violated (upper limit for biomass peak) so too are conditions 1, 3 and 8 violated (primary production too low, biomass peak too early, and phosphorus export too low); however, only 1% of this subclass violates the most critical condition 7 on maximum phosphate level. Excessive primary production always occurs together with an excessive level of phosphate, and about half of this subclass gives the biomass peak to be either in the required interval of time or too late (40%). Excessive variations in the total phosphorus content of the lake are mostly paralleled by a too early biomass peak, but almost never occur in conjunction with an excessive peak value.

Changing condition 7 from the initial 2.5 mg P to 3.0 mg P allowable in the mixed period, decreased the number of violations of this condition from 7201 to 5680, and resulted in 665 'behaviour-giving' character vectors. Thus, 372 of the 1521 vectors located in that 'interval' (compare with Fig. 5) do not violate any other condition. Further change in the allowable phosphate level from 3.0 to 3.5 mg P increased the number of behaviour-giving vectors to 1127, with 4126 residual violations of the constraint condition, indicating a fraction of about 500 potential 'behaviour' vectors within that interval. For comparison, a reduction of the allowable values from the original 2.5 to 2.0 mg m^{-3}, decreased the number of behaviour runs quite dramatically to 68 with a corresponding number of constraint violations of 8565. Again, a considerable number of the vectors in that interval (more than 1000 of the total of 1350) were already violating at least one other constraint condition. In addition, as another example, changing condition 3 from day 60 to day 50, resulted in only two additional 'behaviour' vectors, although the number of violations of condition 3 dropped from 5108 to 5074. The remaining 32 samples thus give a model response that violates at least one other condition.

In conclusion, it can be said that although the specification of some of the constraint conditions is rather critical for the resulting character vector separation, the high degree of coincidence makes the method less sensitive to the individual conditions. This analysis may give some indication of where further efforts in data collection or analysis should be concentrated. Admittedly, however, evaluation of the sensitivity of the approach to the behaviour definition should be carried out in terms of response probability distributions for predictions, which remains to be done.
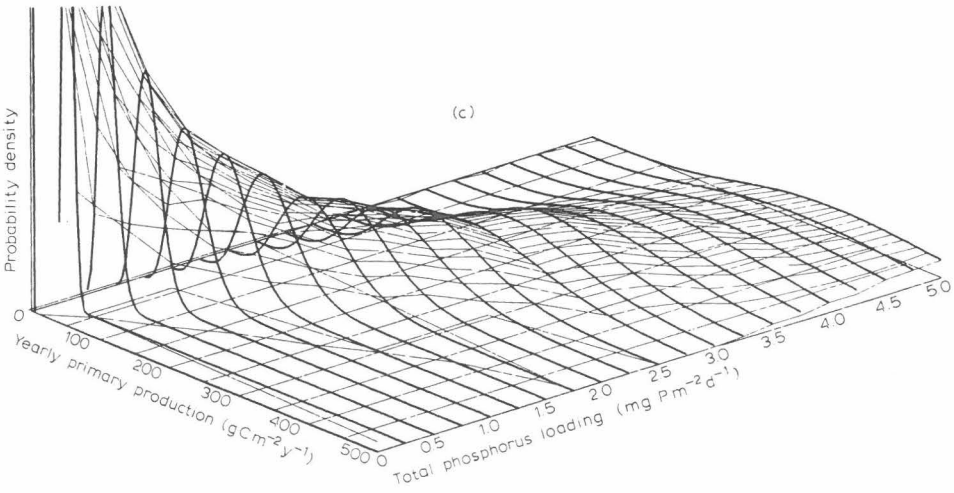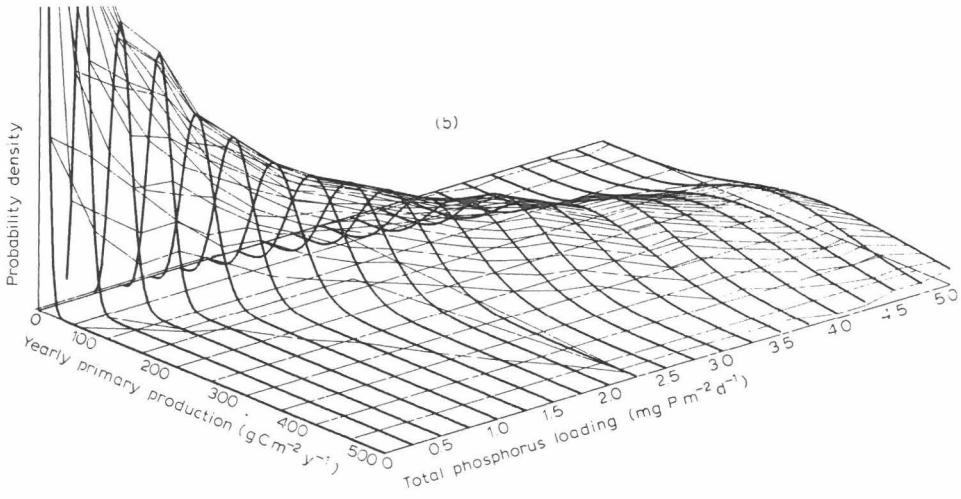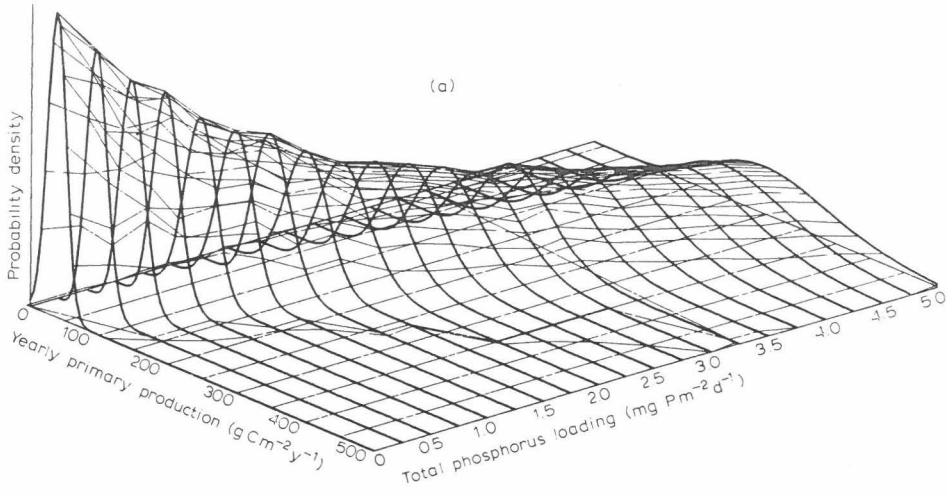
Projections into the future

Having established an ensemble of 'model calibrations' for the range of empirical conditions covered by the behaviour definition, this ensemble can now be used for making predictions of the lake system's response to changes

in nutrient loading. The mean total phosphorus loading in this ensemble was estimated to be $1 \, \text{mg P m}^{-2} \, \text{d}^{-1}$ (S.D. = 0.33), which corresponds well with independent field estimates (Müller, 1979). For the predictions, the loading (character elements 11 and 12 in Table II) were now varied from 0–5 mg P $\text{m}^{-2} \, \text{d}^{-1}$ in steps of 0.25. The proportion of the available phosphorus in the total loading was set to 10% after a series of runs in which ratios of 0, 10, and 25% were compared. For each of the 21 new loading values the first 150 sample character vectors from the behaviour-giving ensemble were taken, thus generating a set of 150 estimates for several output variables (yearly primary production, algae peak biomass, phosphate maximum, phosphorus export, and phosphorus sedimentation) for each loading value for a series of 10 years. Figure 6 summarizes the results for primary production, showing the situation after years 1, 2, 5 and 10.

As indicated in Fig. 6, the probability distributions in the higher-loading classes level out with time. Taking the coefficient of variation as a measure of prediction uncertainty, a saturation curve type pattern in time can be observed for this measure (see also below). When plotting this coefficient of variation against loading for the first year's response, a distinct minimum, in the empirical (observed) range of loadings, can also be observed (Fig. 7). One may conclude therefore, that prediction uncertainty increases with increasing extrapolation, away from both the present time and the presently observed input loading conditions. Certainly, the predictions for larger deviations from the empirical situation are rather trivial after only a few years: an estimate of yearly primary production between 100 and 1000 g C m^{-2} for a phosphorus loading of 5 mg $\text{m}^{-2} \, \text{d}^{-1}$ is certainly of little value as a prediction. However, it should serve as a warning to the analyst that the uncertainty in the data available or the variability of the system itself simply do not support such an extrapolation.

The samples of predictions can now be interpreted in terms of the original problem setting. Only the loading values close to the empirical range result in meaningful distributions in the ten-year projections, but these are of course the most interesting and 'realistic' alternatives to be studied. At the end of the 10-year period of simulations the lake system is found to be in a kind of new dynamic equilibrium with regard to the output variables considered. Whereas the variability of the predictions rapidly increases during the transient stage of the first six years (or less in some cases) after a change in the loading, this variability has stabilized by the end of the simulation period (the somewhat unrealistic zero-loading class is omitted from these evaluations). The time to reach a new equilibrium was found to be related to the relative change in the loading. The comparison of different loadings in terms of primary production, peak biomass, or phosphate level can now be made by either comparing the mean levels and their confidence
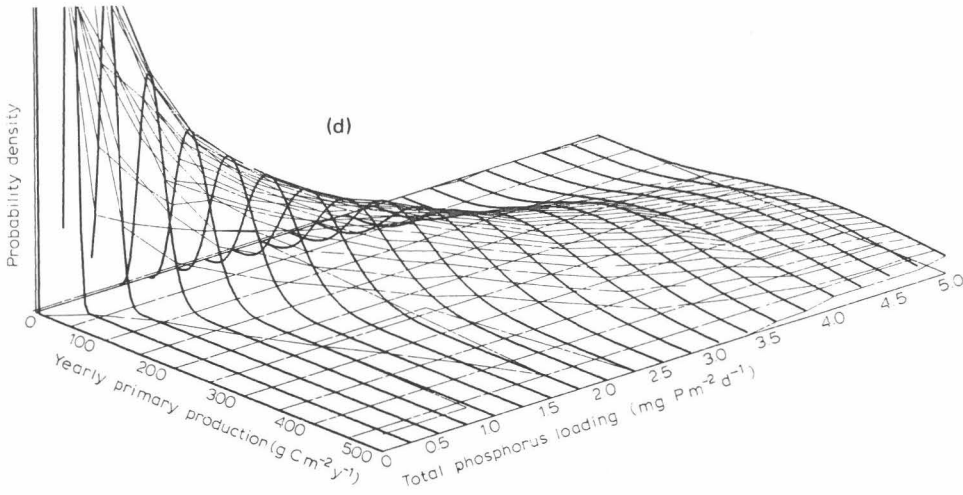
(a)

(b)

(c)

Fig. 6. Estimates of primary production for different loading levels (293 runs each): (a) simulation year 1; (b) simulation year 2; (c) simulation year 5; (d) simulation year 10. Empirical range is used for initial conditions for the first year; simulation is extended for a period of ten years.
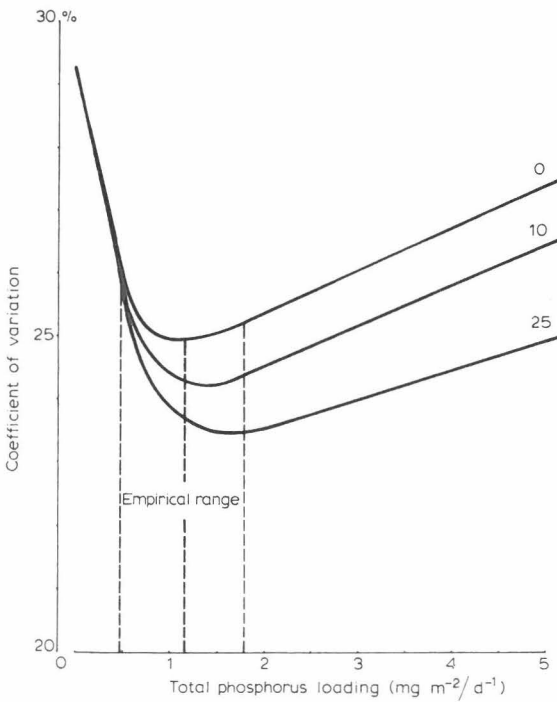

Fig. 7. Coefficient of variation for the model response 'primary production' versus extrapolation in loading; the three different curves represent three percentages of orthophosphate in the total loading.

intervals, or by comparing the probabilities for reaching or exceeding certain levels. This is especially interesting, as in fact almost all of the probability curves fitted are skewed, which clearly implies that a simple comparison of mean values might be misleading for certain problems.

DISCUSSION

This paper, in line with a growing number of other publications (e.g. Beck et al., 1979; Di Toro and van Straten, 1979; O'Neill and Gardner, 1979; Reckhow, 1979; van Straten and de Boer, 1979; Hornberger and Spear, 1980; Spear and Hornberger, 1980), has discussed the problem of uncertainty in ecological modelling. We have, in fact, taken the scope of the paper one step further to include an examination of problems of arbitrariness. In order to carry out an analysis in which these problems of uncertainty and arbitrariness can be considered explicitly we have prepared a method for model 'calibration' and subsequent prediction with the model which is based on the idea of Monte Carlo simulation. Given the uncertainties in the field data and our a priori knowledge of a system's behaviour, 'calibration' of the model depends upon a 'vague' definition of observed behaviour and a comparison with this definition of responses obtained from the model. The structure of the relationships among the variables represented in the model is fixed a priori. The parameter values and other information required to run the model, such as input forcing functions, are gathered together in a character vector, random realizations of which are used to generate a large sample of simulated responses. Each realization of the character vector is drawn from prespecified rectangular distributions for the character-vector elements; each simulation run thereby obtained is classified as either 'giving the defined behaviour' or 'not-giving the defined behaviour'. All the character vector realizations that give the defined behaviour are then regarded as equally good 'calibrations' of the model and they can subsequently be used to generate distributions of the future behaviour of the system under study.

There are five key themes of ecosystems modelling which we wish to discuss in this section: 'uncertainty', 'arbitrariness', 'calibration', 'prediction', and (to a lesser extent) 'management'. Uncertainty, as related to model calibration, is considered explicitly both in defining the range of observed behaviour patterns and in defining the intervals of allowable parameter values and forcing functions (the character vector elements). When the sample of behaviour-giving character vectors is used for generating predictions under changed conditions, the uncertainty of the prior calibration process is preserved and will propagate forward with the predictions. For the case study reported (the Attersee problem) it is clear that the data available for calibration are particularly uncertain. The effects of this are clearly

reflected in the model predictions, which tend quickly to degenerate into rather trivial and meaningless statements about the future. This is, of course, a clear warning that the data available and the observed variability of the system itself simply do not support such extrapolation. We would indeed prefer to make predictions about the range and variability of behaviour patterns rather than 'best' estimates of possibly meaningless average values.

Arbitrariness in ecosystems modelling is at first sight most apparent in the problem of model calibration. There are more or less arbitrary choices that have to be made for the structure of the model, for the specification of the behaviour definition, and for the permissible ranges of the character vector elements. But ultimately, as with the effects of uncertainty, it is for the purposes of prediction that the effects of arbitrariness are important. In this paper, therefore, we have examined the sensitivity of the model response to the arbitrary choices of the calibration procedure. It has not been possible, however, to study the influence of our particular choice of model structure. One could propose that the approach be applied to alternative model structures in parallel, but given a study with only one hypothesis for the model structure, it is necessary to supply the appropriate qualifying statements for any interpretations placed on the prediction results.

There are several challenging aspects of the method adopted here for model calibration. Since the model structure was fixed a priori in this study we have not addressed the difficult problem of model structure identification (e.g. Beck, 1979, 1980); the 'calibration' exercise here has essentially been concerned with a 'parameter estimation' problem. Nevertheless, the induced distributions of the parameter values (in the character vector) that result from the analysis can provide not only meaningful insight into the nature of the system's behaviour but also a posteriori information about the adequacy of the model structure. For instance, a high variability or absence of clustering in the parameter values that give the defined behaviour could indicate an over-parameterization of the model. In other words, this is the equivalent of the problem of parameter identifiability that is commonly encountered in modelling poorly-defined systems (Young, 1978; Young et al., 1978). We might alternatively suggest that such 'random' distributions of the behaviour-giving parameter values are an indication of 'true' stochastic variability in the parameters. However, this seems a less plausible interpretation, and again it is important to recall that the data of the Attersee case study exhibit especially high levels of uncertainty. It is also pertinent to consider in more detail the nature of the criterion for model 'calibration', i.e. the 'vague' definition of the system's behaviour. As knowledge of observed behaviour accumulates and improves so it is possible to imagine that the behaviour constraints can be made both more numerous and narrower. Thus the 'vague' definition becomes progressively less vague. Does this suggest

that the logical conclusion of this argument is that the 'vague' definition eventually approximates a least-squares-like criterion? It does not: the 'vague' definition of system behaviour embodies no such concept as a degree of closeness to a 'true' behaviour; it is merely a binary classification of parameter values that either 'give' or 'do not give' the defined behaviour. Accordingly, one could thus speculate on the merits of a ternary classification procedure with a 'marginal' class separating the behaviour-giving and non-behaviour-giving classes.

The fourth theme for consideration in this discussion is prediction. Of primary concern here is a problem of the persistence of 'extreme' conditions. The problem has its origins in the 'calibration' procedure. In defining the system's behaviour a number of aggregate yearly observations and features were added together. The behaviour definition was of necessity particularly wide in our case study and thus some of the behaviour-giving character vectors reflect conditions of behaviour (towards the boundaries of the defined ranges) which we would intuitively, but not formally, term 'extreme' conditions. When these specific character vector realizations are used for prediction, say for ten years into the future, this is tantamount to assuming that extreme conditions persist for long periods of time. Whether such persistence of extreme conditions contributes to a greater apparent variability of future behaviour is a difficult question to answer. However, it is a question that deserves further study.

Finally, let us turn to the topic of ecological modelling in a management framework. Management usually requires answers now to questions of immediate significance; and these answers must be derived on the basis of the currently available observations. The choice that management has between the obscure, uncertain predictions given here and a detailed, deterministic, and apparently confident statement of future behaviour, is indeed a sensitive issue. What choice would any of us make between an honest, but depressingly ambiguous prediction and a faithfully-promoted, complex, and seemingly accurate prediction? Contributors to a book on adaptive environmental management (Holling, 1978) discuss the 'issue of uncertainty' in considerable depth. The idea of benefiting from uncertainty, in terms of the learning process and in the design of new experimental or monitoring programmes to reduce critical uncertainties, is one with which this paper is in sympathy. Our conclusion is that there should be careful and critical appraisal of 'deterministic' modelling and prediction exercises, which tend to ignore uncertainty and which tend to attract the approval of being accurate because they are complex and contain much detail.

Österreich. The authors gratefully acknowledge R.J. Anderson, Jr. for help-ful suggestions on set notation.

REFERENCES

Beck, M.B., 1979. Model structure identification from experimental data. In: E. Halfon (Editor), Theoretical Systems Ecology. Academic Press, New York, NY, pp. 259–289.

Beck, M.B., 1980. Hard or soft environmental systems? Int. Inst. Appl. Syst. Anal., Laxen-burg, Austria, WP-80-25, 29 pp.

Beck, M.B. and Young, P.C., 1976. Systematic identification of DO-BOD model structure. Proc. Am. Soc. Civ. Eng., J. Environ. Eng. Div., 102: 909–927.

Beck, M.B., Halfon, E. and van Straten, G., 1979. The propagation of errors and uncertainty in forecasting water quality. Int. Inst. Appl. Syst. Anal., Laxenburg, Austria, WP-79-100, 54 pp.

Benson, M., 1979. Parameter fitting in dynamic models. Ecol. Modelling, 6: 97–115.

Di Toro, D.M. and van Straten, G., 1979. Uncertainty in the parameters and predictions of phytoplankton models. Int. Inst. Appl. Syst. Anal., Laxenburg, Austria, WP-79-27, 34 pp.

Fedra, K., 1979. A stochastic approach to model uncertainty: a lake modelling example. Int. Inst. Appl. Syst. Anal., Laxenburg, Austria, WP-79-63, 46 pp.

Fedra, K., 1980. Mathematical modelling—a management tool for aquatic ecosystems? Helgoländer wissenschaftl. Meeresuntersuchungen (in press).

Fedra, K., 1981a. Estimating model prediction accuracy: a stochastic approach to ecosystem modelling. ISEM, 2nd State-of-the-Art Conf. Ecological Modelling, Liège, Belgium (in press).

Fedra, K., 1981b. A Monte Carlo approach to estimation and prediction. In: M.B. Beck and G. van Straten (Editors), Uncertainty and Prediction of Water Quality. Proc. Task Force Meet., Int. Inst. Appl. Syst. Anal., Laxenburg, Austria, (in press).

Flögl, H., 1974. Die Reinhaltung der Salzkammergutseen. Österreichische Wasserwirtschaft, 26: 5–11.

Halfon, E. and Beck, M.B., 1980. Uncertainty in model predictions: a systems approach (in preparation).

Holling, C.S., 1978. (Editor), Adaptive Environmental Assessment and Management. Wiley, Chichester, 377 pp.

Hornberger, G.M. and Spear, R.C., 1980. Eutrophication in Peel Inlet, I. Problem-defining behaviour and a mathematical model for the phosphorus scenario. Water Res., 14: 29–42.

Imboden, D. and Gächter, R., 1978. A dynamic lake model for trophic state prediction. Ecol. Modelling, 4: 77–98.

Jolankai, G. and Szöllősi-Nagy, A., 1978. A simple eutrophication model for the bay of Keszthely, Lake Balaton. In: Modelling the Water Quality of the Hydrological Cycle (Proc. Baden Symp., September 1978), IAHS Publ. No. 125, pp. 137–150.

Jorgensen, S.E., Mejer, H. and Friis, M., 1978. Examination of a lake model. Ecol. Modelling, 4: 253–278.

Lewis, S. and Nir, A., 1978. A study of parameter estimation procedures of a model of lake phosphorus dynamics. Ecol. Modelling, 4: 99–118.

Müller, G., 1979. Jahresbericht, 1978. Arb. Lab. Weyregg, 3/1979, 261 pp.

O'Neill, R.V. and Gardner, R.H., 1979. Sources of uncertainty in ecological models. In: B.P. Zeigler, M.S. Elzas, G.J. Klir and T.I. Ören (Editors), Methodology in Systems Modelling and Simulation. North-Holland, Amsterdam, pp. 447–463.

Reckhow, K.H., 1979. The use of a simple model and uncertainty analysis in lake management. Water Res. Bull., 15: 601–611.

Spear, R.C. and Hornberger, G.M., 1980. Eutrophication in Peel Inlet, II. Identification of critical uncertainties via generalized sensitivity analysis. Water Res., 14: 43–49.

Tiwari, J.L., Hobbie, J.E., Reed, J.P., Stanley, D.W. and Miller, M.C., 1978. Some stochastic differential equation models for an aquatic ecosystem. Ecol. Modelling, 4: 3–27.

Van Straten, G. and de Broer, B., 1979. Sensitivity to uncertainty in a phytoplankton-oxygen model for lowland streams. Int. Inst. Appl. Syst. Anal., Laxenburg, Austria, WP-29-28.

Young, P.C., 1978. General theory of modelling for badly defined systems. In: G.C. Vansteekiste (Editor), Modelling, Identification, and Control in Environmental Systems. North-Holland, Amsterdam, pp. 103–135.

Young, P.C., Hornberger, G.M. and Spear, R.C., 1978. Modelling badly defined systems—some further thoughts. In: Proc. SIMSIG Simulation Conf., Australian National University, Canberra, pp. 24–32.