IDENTIFYING MODELS OF ENVIRONMENTAL SYSTEMS' BEHAVIOUR

M.B. Beck

WP-81-158
December 1981

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

In recent years there has been considerable interest in developing models for river and lake ecological systems, much of it directed towards the development of 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. IIASA's Resources and Environment Area's Task 2 on "Environmental Quality Control and Management" is addressing such concerns, one of its principal themes being to develop a framework for modelling poorly-defined environmental systems.

This paper considers the question of how theories are developed about the behavior of large, complex systems such as those typically relevant to managing environmental quality. It extends earlier discussions of model structure identification (see RR-80-4 and RR-81-4) and provides a more philosophical interpretation of this particular problem and approaches to its solution.

Together with WP-81-108 by L. Somlyódy ("Modelling a Complex Environmental System: The Lake Balaton Study"), this paper has been prepared for a special issue on IIASA's work of the journal "Mathematical Modelling".

ABSTRACT

This paper addresses the question of how theories are developed about the behaviour of large, complex systems such as those typically encountered in managing environmental quality. The specific problem considered is that of model structure identification by reference to experimental, in situ field data. A conceptual definition of this problem is given in terms of the notion of testing model hypotheses to the point of failure. An approach to solving the problem is proposed in which the use of recursive model parameter estimation algorithms is a central feature. This approach is illustrated by a case study in developing a dynamic model of water quality in the Bedford Ouse river in central-eastern England. The results are organized around the two principles of attempting to falsify confident hypotheses and of speculating about relatively uncertain hypotheses in order to modify inadequate prior hypotheses. The essential difficulty demonstrated by the case study is one of absorbing and interpreting the diagnostic evidence of field data analysis and this is ultimately a difficulty associated with the complex and intrinsically indivisible nature of large-scale systems.

TABLE OF CONTENTS

1. INTRODUCTION

According to the spectrum introduced by Karplus [1] environmental systems' analysis lies midway between the two extremes of analyzing socio-economic systems and electrical network analysis. This gives rise to rather special problems in the analysis of environmental and, more specifically, water quality-ecological systems. On the one hand, a priori theory, with its basis in the physical and biological sciences, would seem to be capable of predicting observed behaviour relatively accurately. On the other hand, however, it is especially difficult to conduct planned experiments against which a priori theory can be evaluated. In these somewhat ambivalent circumstances there has arisen a growing incompatability between that which can be simulated in principle with a model and that which can be observed in practice. To a great extent this accounts for the gap that has developed between the "larger" simulation models, with which there is little hope of conducting rigorous calibration exercises given currently available field data, and those much "smaller" models that have been so calibrated.

The specific problem to be considered in this paper is that of model structure identification by reference to experimental, in situ field data. To see why this is a problem, however, it is first necessary to summarise briefly some limitations in a widely accepted approach to water quality-ecological modelling. According to this approach it is generally assumed that one can (conceptually) subdivide the field system into smaller, individual components, whose (conceptual) behaviour can usually be approximated by laboratory-scale replicas (for example, chemostat and open-channel flow experiments). Submodels for these components are assumed to be "verifiable" against experimental observations of the behaviour of the replica; and the model for the field system can be assembled by linking together the sub-models. Thus the content of the model is supported by arguments that admit extrapolations from laboratory systems and equivalent or similar field systems. At the stage of model calibration the tendency is to assume that a priori theory is correct unless demonstrably inadequate. It is especially difficult to demonstrate inadequacy, and the need to question the validity of the original extrapolations is thus all too easily likely to remain obscured.

The argument that the extrapolations inherent in the above approach are legitimate would appear to remain in doubt unless one can develop and apply a complementary approach that provides a more direct evaluation of the prior hypotheses about observed system behaviour, without dividing the system into its component parts. Model structure identification is a fundamental part of that complementary approach: it has to do with the questioning so easily set aside because of the imperfections of the available field data; it is a problem for which seemingly few systematic methods of solution have been developed; and, possibly most significant, it requires a subtle but important change of attitude towards modelling. In spite of very many laboratory-scale experiments and a number of major field studies, current knowledge of the structure of the relationships among the mineral, organic, and microbiological components of an aquatic ecosystem is still quite uncertain. Too much confidence has been placed in a priori theory. Perhaps, in Popper's terms [2], environmental systems have been modelled as though they were "clocks", being "regular, orderly, and highly predictable", whereas they may well be more like the "irregular, disorderly, and more or less unpredictable" "clouds". This reflects simply a change of attitude, because, as evident in Somlyódy's papers [3], [4], there is clearly a spectrum of regularity and orderliness associated with the prior knowledge relevant to water quality-ecological modelling (ranging from hydrodynamics to biology). In short, central to the problem of model structure identification is the question: how are theories developed about the behaviour of large, complex systems given the assumption that observations can be obtained (and subsequently interpreted) from experiments broadly similar to the classical form of experimentation in laboratory science.

The work discussed here, then, on the topic of modelling poorly-defined environmental systems ("poorly-defined" being an expression first used by Young [5]), is part of a Task on Environmental Quality Control and Management within the Resource and Environment Area of IIASA. This essentially methodological component of the Task is complemented by a second theme dealing with case studies in lake eutrophication management, that is, for Lake Balaton, Hungary [6], [7], [8] and for a number of Austrian lake systems [9]. A productive interaction between case-study problem-solving and methodological developments is the cornerstone of the Task's research. In the following, although examples drawn from the lake eutrophication studies would be equally appropriate, such as the results reported by Somlyódy [3], we shall illustrate methodological problems associated with modelling the dynamics of water quality in the Bedford Ouse River (U.K.). This river system in turn provides an informal case-study for the development of a third theme of the Task on operational water quality management [10].

Section 2 of the paper discusses both the problem of model structure identification and an idealised approach to its solution based on the use of recursive parameter estimation. Since model structure identification can be viewed as a matter of iteratively falsifying and speculating about hypotheses, section 3 examines the difficulties of interpreting diagnostic evidence on whether a given model structure (set of hypotheses) is demonstrably inadequate.

2. MODEL STRUCTURE IDENTIFICATION

Usually one associates the exercise of model calibration with curve-fitting and parameter (coefficient) estimation. But the word "calibration" is misleading. It suggests an instrument (here, the model) whose design is complete and whose structure is beyond further argument. All that remains to be done is to make minor adjustments to some of the fittings, i.e. fine-tuning of the parameter values. Calibration of models for water quality-ecological systems, however, is unlikely to be such a simple and straightforward matter. Instead, even before asking the question "Can I estimate the model parameters accurately?", the analyst must first ask himself whether he knows how the variables of the system are related to each other. In particular, one must ask whether information about these relationships can be identified from the in situ field data. Yet most exercises in model calibration have focused solely on the matter of parameter estimation; hence little attention has been paid to the (arguably) more important prior problem of model structure identification.

Let us introduce and qualify a working definition of the problem:

- Model structure identification is concerned with establishing unambiguously, by reference to the in situ field data, how the measured input disturbances, $\underline{u}$, are related to the state variables, $\underline{x}$, and how these latter are in turn related both to themselves and to the measured output responses, $\underline{y}$, of the system under study.

We may note first that this is significantly different from a definition of what may be called model order estimation, a problem in which, for example, the objective is to estimate the orders of the polynomials in an autoregressive/moving-average time-series model (see, for instance, [11], [12], [13]). Second, we may note the importance of the word "unambiguously". A common difficulty in fitting a model to a set of field data is that the error-loss function does not exhibit a well-defined, global minimum. Many combinations of estimates for the model parameter values provide equally good (or bad) descriptions of the observed behaviour; in effect, a uniquely "best" model for the system has not been identified. Such difficulties are often referred to as the problem of identifiability, or the model is said to be over-parametrised and to contain surplus content. This is perhaps a matter of no consequence in terms of fitting the model to the data, but it would certainly have significant implications should the model be used for prediction (as has been argued elsewhere, [14], [15]). One would expect ambiguous statements about future behaviour, although the effects of uncertainty may preclude any conclusion about significant differences among these statements [15].

The essence of the approach to model structure identification, as discussed briefly here and in much greater depth in [5],[15], [16], [17], [18], is based on a restatement of the original problem definition in terms of a parameter estimation problem. Such an approach, however, depends on the availability of an adequate set of time-series field data, a condition which is by no means always satisfied. Even so, for situations of scarce data the development of a roughly parallel approach is apparent in a recent paper by Fedra [19].

In order to outline the approach, albeit in a conceptual sense, let us imagine that the state variables $\underline{x}$ in a model may be represented by the nodes of Figure 1(b) and that the parameters $\underline{\alpha}$ are visualised as the "elastic" connections between the state variables. Without going into details, let us also assume that the parameters of the model can be estimated recursively, i.e. such that estimates $\underline{\hat{\alpha}}(t_k)$ of the parameter values can be obtained for each sampling instant t_k within the sequence of time-series observations (for discussions of recursive estimation, see, for example, [16], [20], [21]).

If now the assumption has been made that all the parameters have values that are constant with time, yet a recursive algorithm yields an estimate of one or more of the parameters that is significantly time-varying, one may question the correctness of the chosen model structure. We can argue this point as follows. The general tendency of an estimation procedure is to provide estimates $\underline{\hat{x}}$ of the state vector, or some functions thereof, i.e. $\underline{\hat{y}}$, that track the observations $\underline{y}$. Hence, if any <u>persistent structural discrepancy</u> is detected between the model and "reality" (in other words, the errors $\underline{\varepsilon} = (\underline{y} - \hat{\underline{y}})$ exhibit a significantly non-random pattern), this will be revealed in terms of <u>significant adaptation of the estimated parameter values</u>. There may well be good reasons for why the parameter estimates vary with time, and, indeed, that is precisely what one is looking for.

Starting with Period 1 of Figure 1(a), however, let us continue to sketch the outline of the approach. The model responses ($\hat{\underline{y}}$) and output observations ($\underline{y}$) are essentially in agreement over this period and there is no significant adaptation of the parameter estimates (according to Figure 1(c)). At the beginning of Period 2, however, there is a persistent discrepancy between $\hat{\underline{y}}$ and $\underline{y}$. It might be supposed, for example, that the underlying cause of the discrepancy is an inadequacy in the behavior simulated for x_1 and x_2, that α_1 is sensitive to this discrepancy (Figure 1(b)), and that (persistent) adaptation of the estimate $\hat{\alpha}_1$ (Figure 1(c)) partly compensates for the error between $\hat{\underline{y}}$ and $\underline{y}$. Again in the third period there is disagreement between the observations and model responses, which leads to adaptation of the estimate $\hat{\alpha}_2$.

The example of Figure 1 is clearly an ideal view of how a recursive estimation algorithm should be employed for model structure identification. In fact it is an idealised framework developed largely, but not entirely, from a particular case-study in modelling the dynamics of water quality in the River Cam, U.K. [17], [22]. Generalisation from a single example is undoubtedly not without dangers and certainly the results to follow challenge the usefulness of this ideal view. Nevertheless, cast in this particular fashion such an approach has intuitively appealing interpretations. First, and by analogy with the analysis of physical structures, the aim is to expose inadequacy in terms of the "plastic deformation" (Figure 1(c)) of the model structure. Second, and of deeper significance, testing the model structure to the point of failure, that is, the failure of one or more hypotheses, can be said to be consistent with Popper's view of the scientific method [23]. And Popper's view of the scientific method is in turn exercising a growing influence over the discussion of modelling the behaviour of environmental and similar systems [5], [19], [24], [25], [26].

Especially pertinent here is Holling's remark that "...the model is [to be] subjected to a range of tests and comparisons designed to reveal where it fails" [24]. This, with emphasis on the words "range" and "designed to reveal" sets a suitable guiding principle for solving the problem of model structure identification. But to have revealed that the model structure is inadequate is merely a part of the solution, and actually a <u>relatively</u> easy part. If we extend the example of Figure 1 one further step, let us suppose that the first (model) hypothesis has been identified as failing, according to Figure 2(a). Now assume that a second hypothesis can be generated in some way--which is a complementary part of the solution--and that it has the structure of Figure 2(b) with an additional state variable (x_5) and two new parameters (α_5, α_6). It may well be that calibration of the second model against the field data yields effectively invariant parameter estimates and hence the analyst can accept the adequacy of this model structure as a conditionally good working hypothesis.

The basic aim of model structure identification is thus to seek plausible hypotheses for apparently "unexplained" relationships in a set of field data. The approach outlined above exploits the idea of curve-fitting as a "means-to-an-end" and not as an "end" in itself. Falsifying the model structure, or components thereof, rests partly upon judgements about absurd parameter values, or about implausible variations in the parameter values. Unless these variations and values can be defended by logical argument, then it must be conceded that the structure of the model does not match the structure underlying the observed patterns of behaviour.

It would be wrong, however, to assume, because of the exclusive discussion of an approach based on recursive parameter estimation, that this approach is a panacea. The benefits to be derived from a range of procedures have already been emphasized and are apparent in the Cam case study [22]. This is only one approach applicable to a certain sector of the overall problem defined for a restricted set of conditions; yet it is an approach that has yielded considerable insight into the nature of the problem.

In the following we shall focus on two types of critical difficulties in applying the above approach to model structure identification, that is: the difficulty of revealing that a hypothesis is absurd, which is really the most demonstrable form of inadequacy; and the difficulty of synthesizing the diagnostic evidence in order to speculate about how to modify an inadequate prior hypothesis. Our purpose is to expose weaknesses and limitations both in the technical effectiveness of recursive parameter estimation as a method of solution and, more fundamentally, in the appropriateness of the approach. As with model structure identification, so too with the approach itself, establishing what is wrong or inadequate is the key to improvement and progress.

3. DESIGN FOR FAILURE AND SPECULATION

If solving the problem of model structure identification depends strongly upon revealing absurd hypotheses, an easily recognizable difficulty is that in situ field data subject to high levels of uncertainty are hardly likely to yield such revelations. There are, however, more subtle aspects of the nature of field data from environmental systems that place equally, if not more, awkward constraints on the likelihood of success in model structure identification. The patterns of time-series observations typically available for analysis reflect experiments--if indeed they can be so called--that are successively less good approximations of the classical, planned experiments of laboratory science [15]. In all but a few cases the observed perturbations in system behaviour do not conform with the desirable attributes of data usually expected for the identification of models for, for example, aircraft and industrial process control [12], [27]. And since it is in areas such as these latter that many of the methods of analysis have originated [28], recursive estimation included, one finds that there is an impressive array of techniques that perform well on well-posed problems, yet a dearth of techniques that can perform adequately on the ill-posed problems of environmental systems analysis.

It is tempting to blame a lack of success on poor data and inappropriate analytical methods. But this would be misleading and, in any case, current constraints are not destined to persist into the future. Consider, for example, the ever-growing potential for generating data from environmental monitoring networks and consider also the principal asset of a recursive estimation algorithm, that is, to generate model parameter estimates at each instant of time t_k in a time-series. There is every possibility that future critical constraints will be dominated by the inability to absorb and interpret the diagnostic evidence of data analysis. In fact, these constraints are ultimately a function of the complexity and indivisibility of large-scale systems. It is to the difficulties of conducting an analysis in the face of such problems that we now turn.

From the generalisation of the River Cam case study, to which passing reference has been made earlier, it is possible to propose a tentatively broader organising principle for the procedure of model structure identification. Hence, let us simply suggest that the analyst is concerned with conducting experiments (in a loose sense) on and with the model structure, where these experiments can have the following two distinctly different orientations (or objectives):

(i) in the process of falsifying a given model structure;

(ii) in the process of (creative) speculation about alternative hypotheses.

These two processes are probably best viewed as mutually exclusive, for reasons we shall discuss later, and, quite appropriately, they reflect the two-step nature of solving the problem.

The case of the Bedford Ouse River in central-eastern England is a natural extension of the Cam study. From 1972 to 1975 the Department of the Environment in the United Kingdom and the Anglian Water Authority jointly funded a major study of the Bedford Ouse river system in order to evaluate the effects of developing a new city (Milton Keynes) in the upper part of the catchment [20]. It is in the light of tackling this substantially more complex problem of field data anlaysis that we shall be able both to judge the usefulness of the above organising principle and to illustrate the difficulties of interpreting the diagnostic evidence of analysis.

3.1 Failure of the Model Structure

Let us look first at the notion of testing the model structure to the point of failure, that is, the process of falsifying a given set of hypotheses. For the Bedford Ouse example the model structure to be evaluated contains various confident assumptions about the transport and dispersive properties of the river, reaeration, the decay of waste organic matter, and the growth, death, and photosynthetic properties of a population of phyto-plankton. That these should be "confident" assumptions, which has a quantitative counterpart in the specification of the a priori error statistics associated with the model, is an important point. Given the conceptual outline for model structure identification (see also Figures 1 and 2) this is a very deliberate tactic of stressing a relatively rigid structure so that the probability of detecting a significant failure is maximised. In this step of the analysis it would not appear to be particularly useful to express little confidence, a priori, in the model and then to try and identify unambiguously where failure occurs. In such a case the postulated model structure is, as it were, too flexible. Adaptation of the parameter estimates may, or may not, be significant, because one has little confidence in the model, and clear-cut answers cannot be obtained because, in effect, clear-cut questions are not being asked. Flexibility would be more of an advantage at the stage of creative speculation and this is why separation of the two steps is desirable.

Altogether six parameters are to be estimated in identical model structures for the behaviour of interactions between dissolved oxygen (DO), biochemical oxygen demand (BOD, a measure of degradable organic matter), chlorophyll-a (as a measure of phytoplankton populations), and suspended solids concentrations in each of the three reaches of the river system (a total, therefore, of 12 state variables and 18 parameters). Figure 3 shows the recursive estimates of these six parameters for the third (downstream) reach of river. Comparing Figure 3 with the enviable idealised simplicity of Figure 1, one would have great difficulty in answering the question "at what point does the model structure fail?" without even asking the question why it might have failed. The results are a peculiar mixture of both insufficient and redundant hypotheses in the model structure--of, at the same time, under- and over-parameterisation. The considerable non-stationarity of the parameter estimates clearly indicates that the model structure is inadequate. Yet the similar patterns of variability among the different parameters is a symptom of surplus content in the model, i.e. one inadequacy compensates for another. In other words, certain critical features of the structure of the relationships underlying the field data are not included in the model, while no single parameter estimate unambiguously compensates for the obvious inadequacy.

There are apparently some absurd hypotheses. For instance, the recursive estimates of both the maximum specific growth-rate (nonlinear Monod kinetics) and first-order, death-rate constants for the phytoplankton population (Figures 3(b) and 3(e) respectively) become negatively valued. One could argue, as a result, that the former is barely significantly different from zero and that the latter--a linear, negative, death-rate--is perhaps evidence of a preferred linear growth-rate function for the phytoplankton (at least for all but the initial period of the data). But the analyst would be hard pressed to attach great confidence to such conclusions. On balance it might be more appropriate to conclude that the algal population is in a state of equilibrium with neither of the rates of growth and death being independently identifiable from the data.

The principal issue raised by the results of Figure 3 is one of misplaced confidence in a priori theory. It has a specific aspect associated with these results and a more general aspect relating to the introductory comments of the paper. Thus, for example, the remarkable stationarity of the recursive estimate for the reaeration rate constant (Figure 3(a)) is a function of having assumed relatively more a priori confidence in this particular parameter. In other words, the analyst has assumed that if the model is to fail it is unlikely to be a function of an inadequate description of the reaeration process, a point to which we return later. This might be a reasonable assumption since, together with the assumption concerning BOD decay, about which similar questions will be raised shortly, it is a basic component of the <u>classical</u> studies (conducted in 1925) of Streeter and Phelps [30] on river pollution and self-purification. That these assumptions have been used for a long time creates a resistance to challenging their validity. Yet there are good reasons, as demonstrated elsewhere [15], for arguing that the classical assumptions of Streeter and Phelps, and the equally classical assumptions of dispersion in flowing media, represent patterns of behaviour that are not identifiable from this particular set of in situ field data. In this case the problem of identifiability arises because other dominant modes of behaviour--here, especially in the first and second reaches of the river, the growth of a phytoplankton population--almost entirely obscure these less significant modes of behaviour. In a sense, therefore, the assumptions of Streeter and Phelps are, for this example, not testable propositions, and their inclusion in any subsequent model structure is tantamount to an act of faith.

It seems important in a more general sense, therefore, to question the motives for maintaining hypotheses that are not, strictly speaking, falsifiable. The reluctance to set aside convention is strong indeed, and Figure 3(c) illustrates well the conflict that can occur--Young [31] has put forward a cogent and challenging argument on the same point. Given prior experience that the hypothesis of BOD decay is probably not identifiable, a BOD decay rate constant is still retained in the model structure, but with an a priori estimate of zero (day^{-1}). It would be difficult to argue from Figure 3(c) that the subsequent pattern of the recursive estimates prompts the assumption of a significantly non-zero value for this parameter. The problem can thus be summarised as follows. The results of Figure 3 are founded upon the premise that:

(a) "We have confidence in the hypotheses of Streeter and Phelps, but consider current hypotheses about mechanisms of phytoplankton growth as highly speculative."

Such a premise could be reoriented to either of:

(b) "We are confident about our hypotheses for phytoplankton growth, but consider the assumptions of Streeter and Phelps to be highly speculative;"

(c) "All hypotheses are equally speculative."

Perhaps one should cling to the first premise and not reject convention until it is demonstrably inadequate. The obvious dilemma is that just such a clutching at convention, especially in the context of water quality-ecological modeling, may preclude the possibility of revealing inadequacy. And the shift in emphasis as to where greater confidence is placed, from premise (a) through (b) to premise (c), is a specific interpretation of the change in attitude towards modelling discussed in the introduction to the paper.

3.2 Creative Speculation

The process of speculation can be illustrated with results drawn likewise from another part of the Bedford Ouse analysis. It is again assumed (implicitly) that premise (a) above is reasonable so that speculation can be conducted in terms of a vector of lumped parameters representing all the other mechanisms of behaviour (in this case, sources and sinks of DO, BOD, and chlorophyll-a) that are considered to be speculative assumptions. The objective then is to generate plausible hypotheses about why the estimates for these lumped parameters exhibit variations with time (or space), if that is so; to formalise these hypotheses; and to proceed to a subsequent step in the process of falsifying the revised model structure. For the three reaches of the Bedford Ouse system, part of the diagnostic evidence from analysis of this speculation is gathered together in Figures 4 and 5. One could tentatively conclude from these recursive estimates that:

(i) The rate of addition of chlorophyll-a to the system reaches a maximum first (in time) in the third (downstream) reach, then in the second, and lastly in the first (upstream) reach, Figure 4;

(ii) The rate of addition of dissolved oxygen to the first reach is roughly proportional to the observed concentration of chlorophyll-a at the downstream boundary of that reach, Figure 5(a); the rate of addition of dissolved oxygen to the second reach is roughly proportional to the observed concentration of chlorophyll-a, except over the middle period of the record, Figure 5(b); the rate of addition of dissolved oxygen to the third reach is not obviously proportional to the observed chlorophyll-a concentration for most of the time, Figure 5(c).

It would certainly be a bold and imaginative hypothesis that could be synthesised from such evidence and hence lead to the restructuring of the model for the purposes of again attempting to falsify the revised hypotheses. And this is actually a relatively simple example, when compared with the complexity of models frequently discussed in the literature. We have presented the evidence of Figures 4 and 5 primarily so that one can ask the rhetorical question: how would the analyst absorb and interpret this relative wealth of diagnostics? As earlier, to have drawn the possible conclusion that the model fits the data subject to arbitrary variations in one or more of the parameters (as typified by the recursive estimates of Figures 3, 4, and 5) is of no consequence. Rather, it is the process of speculating about why such variations occur that should be highly valued.

3.3 Reconstructing the Experiments of Laboratory Science

In introducing the problem of model structure identification it was assumed that observations could be obtained (and subsequently interpreted) from experiments broadly similar to the classical form of experimentation in laboratory science. We shall further assume that a laboratory experiment is usually designed to test the relationship between, say, two variables (cause and effect) while all other variables associated with the system are maintained at steady, constant values. Clearly the field data available from environmental systems reflect quite imperfect experiments. Let us suppose, nevertheless, that model structure identification is a procedure for reconstructing in situ "experiments" from observed data by (mathematical) analytical methods. In other words it seems reasonable to attempt to design the analysis of model structure identification such that it compensates for the unsteady and extraneous disturbances originating from the "environmental conditions" of the laboratory-type "experiment". An apt example is premise (a) associated with the Bedford Ouse analysis in section 3.1 where the "experiment" would be concerned with identifying the mechanisms of phytoplankton growth and the Streeter-Phelps assumptions would be absorbed into the analytical compensation for the "experimental environment". Another apt example is given in Somlyody's paper [3], where the "experiment" is to identify the relationship between wind stress at a point on the surface of Lake Balaton and the distribution of suspended solids measured in the vertical water column below that point. All other phenomena affecting the vertical distribution of solids, that is, other than sedimentation and the wind-induced resuspension of particles from the bed of the lake, are assumed to be included in the "environmental conditions". This latter would include, for example, solids transported horizontally into the vertical water column and that fraction of the observed suspended solids concentration due to living organic matter, such as a phytoplankton population. In fact, the model for this "experiment", as defined, is sufficiently well posed that the analysis might more fruitfully be "inverted" in order to identify better the relationships assumed in the given definition of the "environment".

In either of the two examples quoted, the skill of the analyst would lie in arranging the analysis such that extraneous interference with the analysis could be filtered out. At first sight this is perhaps a rather attractive view of the true purpose of system identification and time-series analysis. But it presupposes, of course, that that part of the model required to compensate for the experimental "environment" is known a priori with sufficient confidence to permit the full power of the analysis to be directed towards identification of the relationships defined as the "experiment". Such assumptions themselves have to be evaluated. The distinction between what is "known well" and what is "speculation" thus becomes vanishingly small. It is unlikely, as with premise (c) in section 3.1, that all prior hypotheses are equally speculative; rather, a spectrum of degrees of confidence is probable. The freedom to manipulate where greater prior confidence should be placed, however, can thus be seen to be both an advantage and a disadvantage. In its worst form it allows the possibility of prejudicing the diagnosis of failure, as apparent with the results of Figure 3. It is difficult to claim, however tempting it may be, that there is just one "experiment" and its complementary "environment". Instead, it is only possible to state that a number of more or less significant "experiments" are proceeding in parallel. This does not mean that partially isolated experiments cannot be conducted on large-scale field systems--the study of wind-induced resuspension of lake sediments in Somlyody's paper [3] typifies what is possible in this respect. But it does mean that if the analyst aspires to the development of a model for the field system as a whole, then his analysis of the data will have to contend with the intrinsically indivisible character of the system's behaviour.

4. CONCLUSIONS

Many contemporary exercises in water quality-ecological modelling have been conducted without serious consideration of the significance of calibration. It is not an empty appendix to the mainstream developments in water quality modelling. It may only be considered so if one chooses to attach great confidence to a priori theory, thereby renouncing, in effect, much of the questioning that should accompany calibration.

The "questioning" process of model calibration, to which considerable importance is attached, is what has been called here the problem of model structure identification. The procedure proposed for solving (in part) this problem has two primary features: (a) the use of recursive parameter estimation algorithms for the analysis of time-series field data; and (b) the alternate objectives of examining the model structure from the point of view of either falsifying confident hypotheses or creatively speculating about uncertain hypotheses.

The paper has illustrated this approach to model structure identification with a case-study of the Bedford Ouse river system. The relative complexity of this study defines it as what might be called a second-generation study in model structure identification; indeed it raises more questions than it answers. In particular, the Bedford Ouse example challenges the usefulness of the procedure outlined for model structure identification and draws attention to the crucial difficulty of focusing and interpreting the diagnostic evidence of analysis. This example also illustrates the problem of distinguishing between which are confident and which are speculative prior hypotheses, a distinction that is important for implementing the proposed approach. Finally, consideration of an analogy with the planned experimentation of laboratory science, although superficially attractive as an interpretation of model structure identification, leads to the conclusion that the analyst has to contend with the multiplicity of "experiments" inherent in a set of field data from an environmental system. Clearly, complexity, an intrinsically indivisible nature, and not merely uncertainty, are inescapable problems in modelling such large-scale systems.

ACKNOWLEDGEMENT

The author is grateful to the Anglian Water Authority and the (U.K.) Department of the Environment for permission to use the field data from the Bedford Ouse Study. I am also indebted to George Hornberger, Gerrit van Straten, and Peter Young for their constructive, critical reviews of an earlier draft of this paper.

REFERENCES

1. W.J. Karplus, The future of mathematical models of water resources systems. In: G.C. Vansteenkiste (Editor), System Simulation in Water Resources, North-Holland, Amsterdam, pp. 11-18 (1976).
2. K.R. Popper, Of clouds and clocks. In: Objective Knowledge, An Evolutionary Approach, Oxford University Press, Oxford, pp. 206-255 (1972).
3. L.Somlyody, Modelling a complex environmental system: the Lake Balaton case study (this issue).
4. L. Somlyody, Water quality modelling: a comparison of transport-oriented and biochemistry-oriented approaches, Working Paper, WP-81-117, International Institute for Applied Systems Analysis, Laxenburg, Austria (1981).
5. P.C. Young, General Theory of Modeling for Badly Defined Systems. In: G.C. Vansteenkiste (Editor), Modeling, Identification, and Control in Environmental Systems, North-Holland, Amsterdam, pp. 103-135 (1978).

6. G. van Straten, G. Jolankai, and S. Herodek, Review and Evaluation of Research on the Eutrophication of Lake Balaton--A Background Report for Modeling, Collaborative Paper, CP-79-13, International Institute for Applied Systems Analysis, Laxenburg, Austria (1979).

7. G. van Straten and L. Somlyòdy, Lake Balaton Eutrophication Study: Present Status and Future Program, Working Paper, WP-80-187, International Institute for Applied Systems Analysis, Laxenburg, Austria (1980).

8. A.V. Leonov, Mathematical modelling of phosphorus transformation in the Lake Balaton ecosystem, Working Paper, WP-80-149, International Institute for Applied Systems Analysis, Laxenburg, Austria (1980).

9. K. Fedra, Austrian Lake Ecosystems Case Study: Achievements, problems, and outlook after the first year of research, Collaborative Paper, CP-80-41, International Institute for Applied Systems Analysis, Laxenburg, Austria (1980).

10. M.B. Beck, Beyond Planning and design: operational water quality management, Executive Report 7, International Institute for Applied Systems Analysis, Laxenburg, Austria (1981).

11. G.E.P. Box and G.M. Jenkins, Time-series Analysis, Forecasting and Control, Holden-Day, San Francisco (1970).

12. K.J. Åström and P. Eykhoff, System Identification - A Survey, Automatica 7, 123-162 (1971).

13. T. Söderström, On Model Structure Testing in System Identification, Int. J. Control 26, 1-18 (1977).

14. M.B. Beck, Hard or Soft Environmental Systems?, Ecological Modeling 11, 233-251 (1981).

15. M.B. Beck, Uncertainty, system identification and the prediction of water quality. In: M.B. Beck and G. van Straten (Editors), Uncertainty and Forecasting of Water Quality, Pergamon, Oxford (in press).

16. P.C. Young, A Recursive Approach to Time-series Analysis, Bulletin of the Institute of Mathematics and its Applications 10, 209-224 (1974).

17. M.B. Beck and P.C. Young, Systematic Identification of DO-BOD Model Structure, Proc. Am. Soc. Civ. Engrs., J. Env. Eng. Div. 102 (EE5), 909-927 (1976).

18. M.B. Beck, Model Structure Identification from Experimental Data. In: E. Halfon (editor), Theoretical Systems Ecology, Academic Press, New York, pp. 259-289 (1979).

19. K. Fedra, Hypothesis testing by simulation: an environmental example, Working Paper, WP-81-74, International Institute for Applied Systems Analysis, Laxenburg, Austria (1981).

20. A.H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York (1970).

21. A. Gelb (Editor), Applied Optimal Estimation, M.I.T. Press, Cambridge, Massachusetts (1974).

22. M.B. Beck, Random Signal Analysis in an Environmental Sciences Problem, Applied Mathematical Modelling 2, 1, 23-29 (1978).

23. K.R. Popper, The Logic of Scientific Discovery, Hutchinson, London (1959).

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

25. J.M. Maciejowski, Model Discrimination Using an Algorithmic Information Criterion, Automatica 15, 579-593 (1979).

26. J.M. Maciejowski, A Least-Genericity Principle for Model Selection, Report No. 91, Control Theory Centre, University of Warwick, U.K. (1980).

27. I. Gustavsson, Survey of applications of identification in chemical and physical processes, Automatica, 11, 3-24 (1975).

28. P. Eykhoff, System Identification - Parameter and State Estimation, Wiley, Chichester (1974).

29. Bedford Ouse Study, Final Report, Anglian Water Authority, Huntingdon, U.K. (1979).

30. H.W. Streeter and E.B. Phelps, A Study of the Pollution and Natural Purification of the Ohio River, Bulletin No. 146, U.S. Public Health Service, Washington, D.C. (1925).

31. P.C. Young, The Validity and Credibility of Models for Badly Defined Systems. In: M.B. Beck and G. van Straten (Editors), Uncertainty and Forecasting of Water Quality,

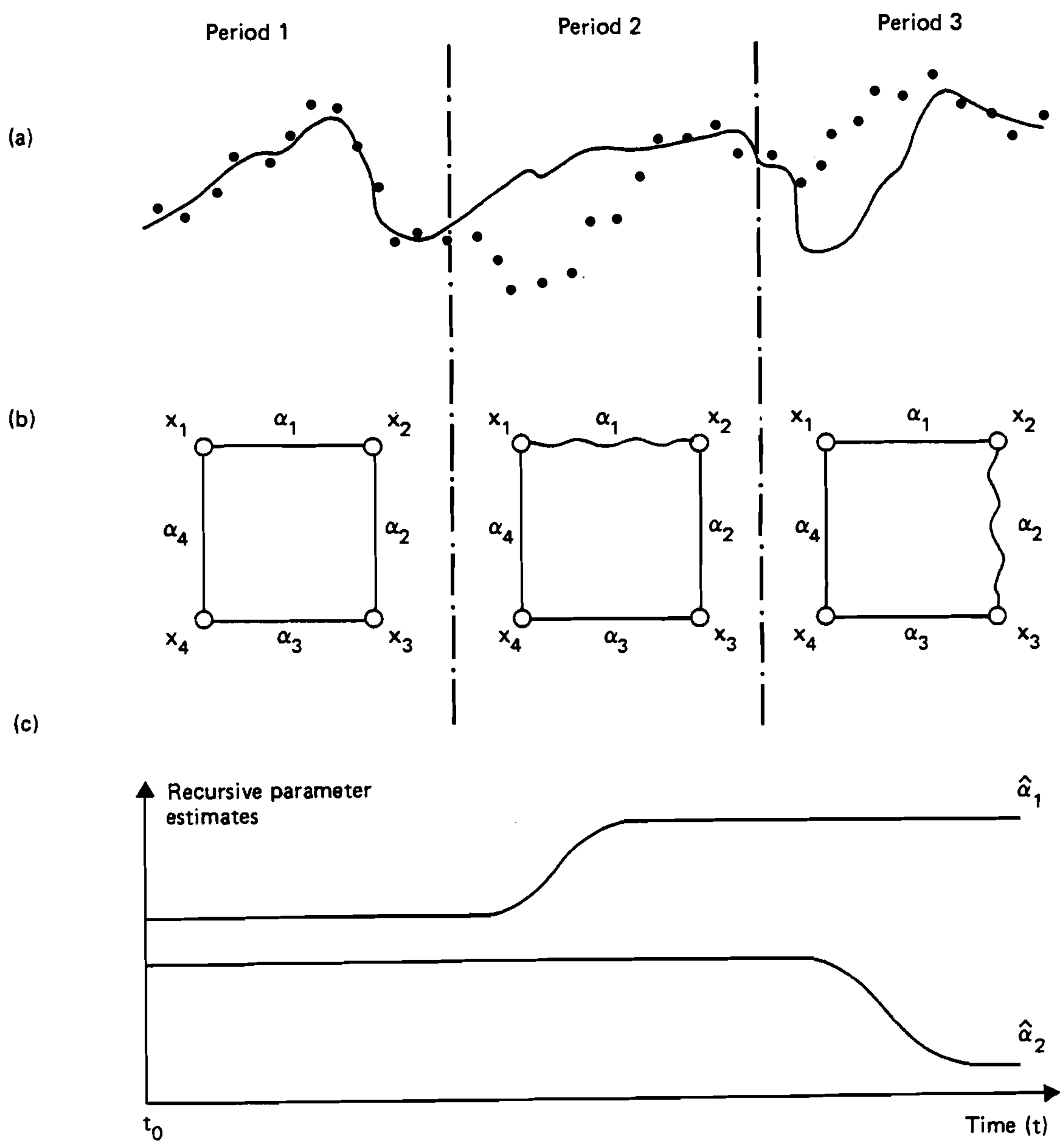

Figure 1. An Illustrative Example showing the Concept of using a Recursive Parameter Estimator in the Context of Model Structure Identification: (a) Hypothetical Model Response and Observations (dots); (b) Conceptual Picture of Model Structure; (c) Recursive Parameter Estimates

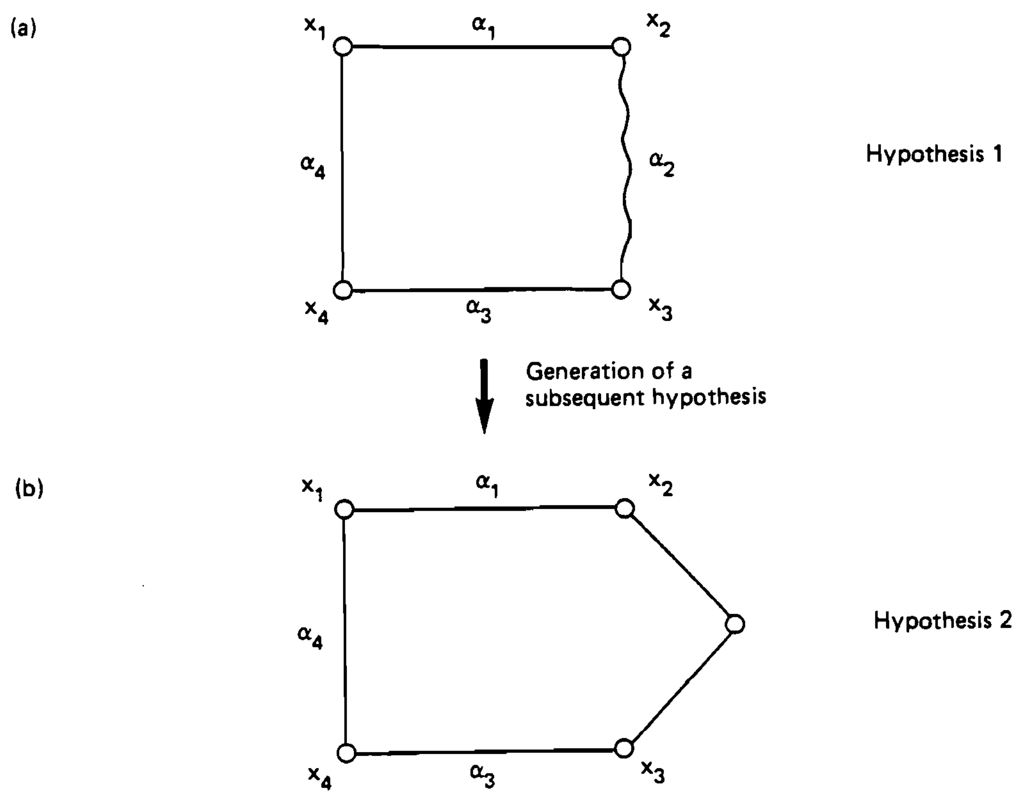

Figure 2. The Process of Model Structure Identification: Revision of the Model Structure and Re-estimation of the associated Parameters (b) on the basis of diagnosing how the prior Model Structure fails (a).

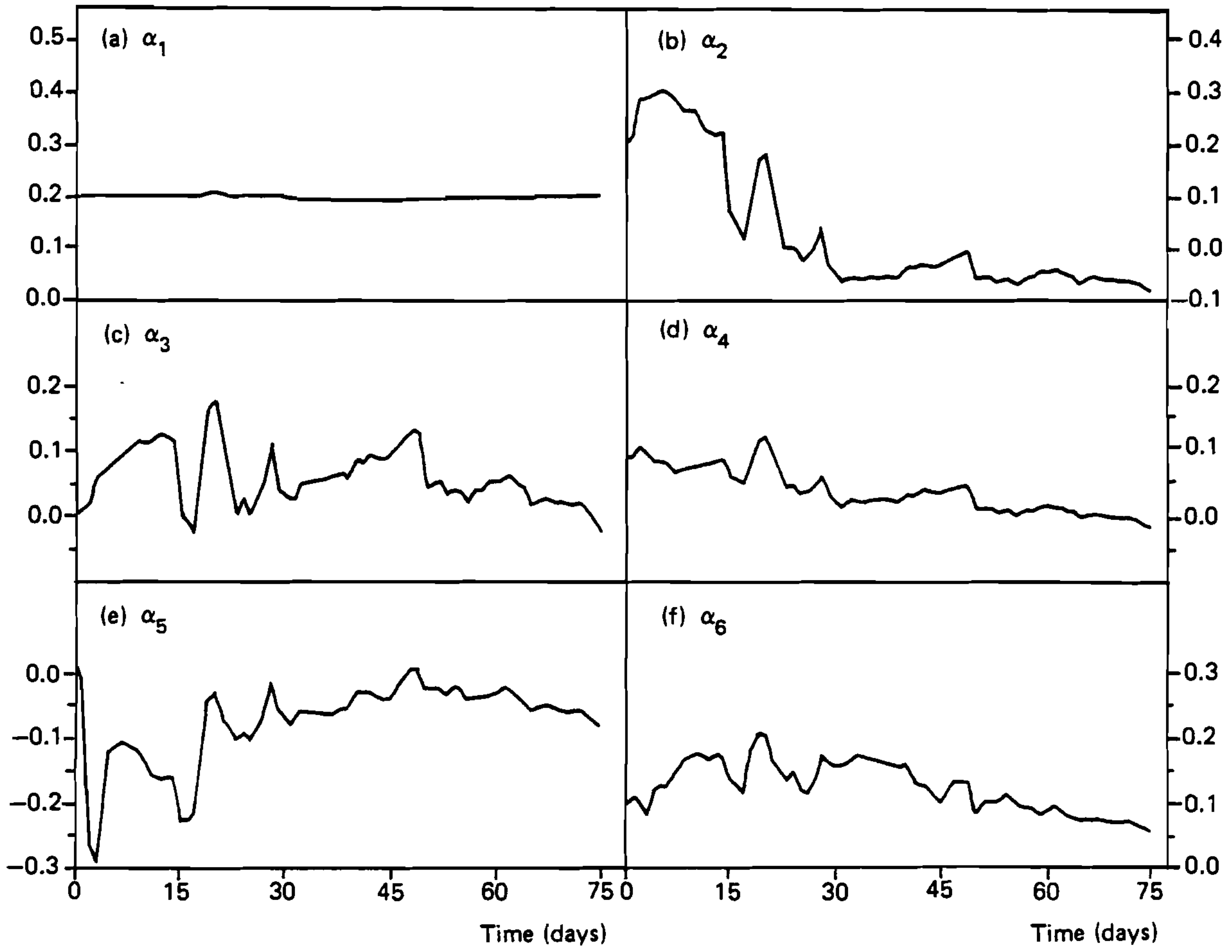

Figure 3. Model Structure Identification (the process of falsifying confident hypotheses) in the Bedford Ouse Case Study (third reach): (a) Reaeration Rate Constant (day^{-1}); (b) Maximum Specific Growth-Rate Constant for Phytoplankton (day^{-1}); (c) BOD Decay Rate Coefficient (day^{-1}); (d) Rate Constant for addition of BOD to reach from Suspended Solid Matter (day^{-1} [gm^{-3} BOD] [gm^{-3} SS]$^{-1}$); (e) Death-rate Constant for Phytoplankton (day^{-1}); (f) Rate Constant for "Loss" of Suspended Soils from the Reach (day^{-1}).

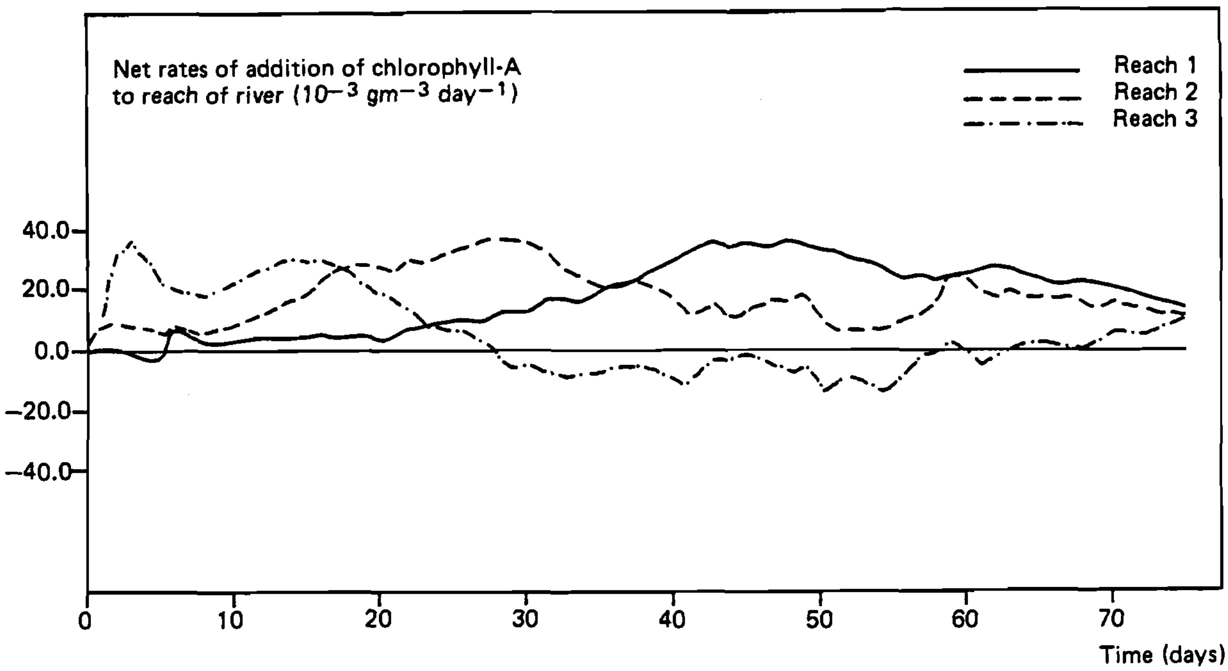

Figure 4. Model Structure Identification (the process of speculation) in the Bedford Ouse Case Study: Recursive Estimates for the Net Rates of Addition of Chlorophyll-A to each Reach of the System

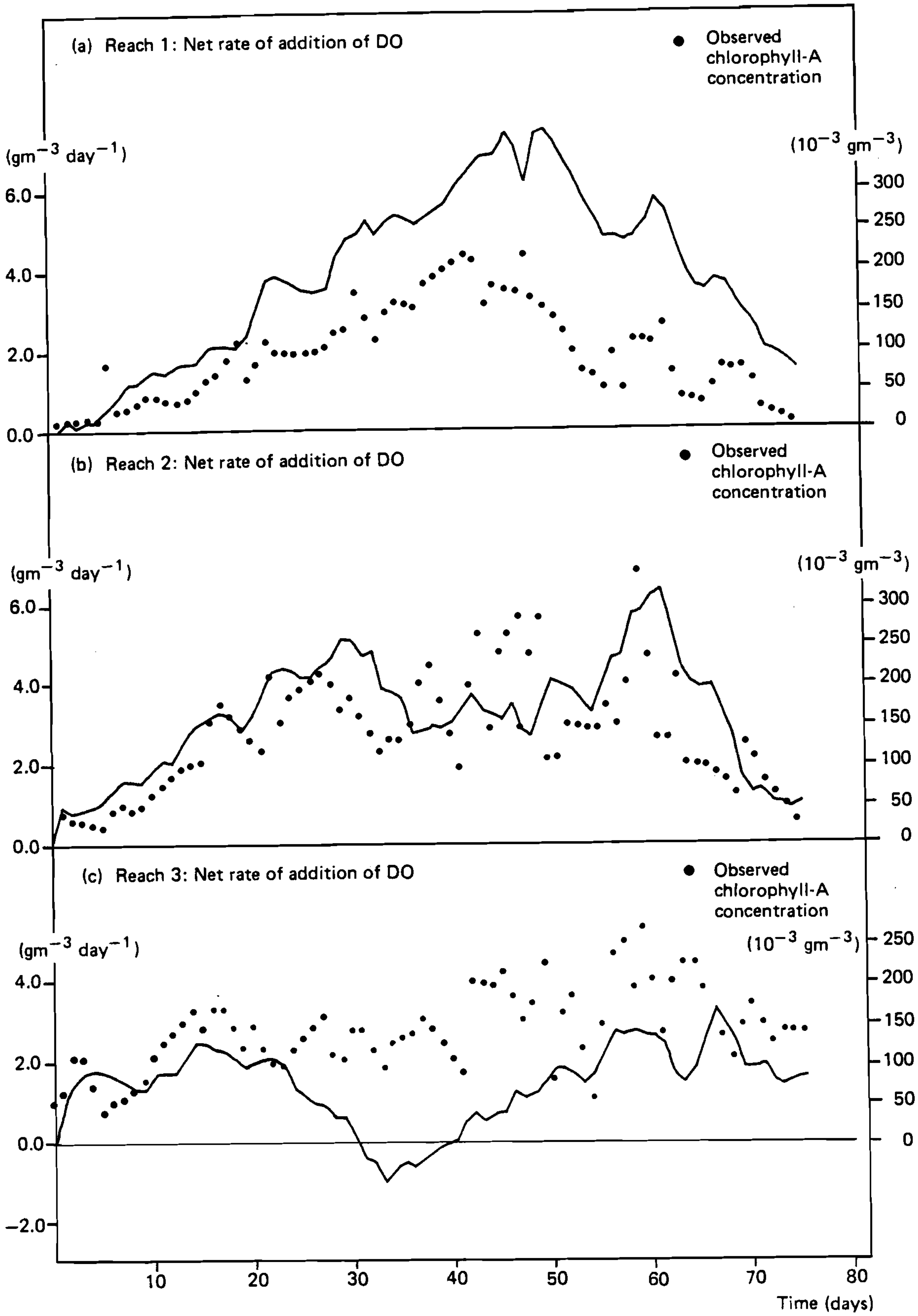

Figure 5. Model Structure Identification (the process of speculation) in the Bedford Ouse Case Study; Comparison of Recursive Estimates for the Net Rates of Addition of DO to each Reach of the System with the observed Chlorophyll-A Concentrations at the Downstream Boundary of each respective Reach.