INSTITUTE
OF
HYDROLOGY

WATER QUALITY MODELLING, FORECASTING AND CONTROL

Proceedings of an International Workshop
at the Institute of Hydrology, Wallingford.

Edited by

P G WHITEHEAD & P E O'CONNELL

PREFACE

In July 1971 a workshop on Real Time Hydrological Forecasting and Control was held at IH in which a small group of specialists worked together on a number of research topics (O'Connell, 1981). Problem areas were identified and research conducted for one month using data and computing facilities available at the Institute.

In 1981 a similar workshop was held but this time the central theme was water quality. A six-week workshop was held during July and August 1981 with the overall objectives of investigating methodological problems, applying water quality modelling and analysis techniques to some real-world problems and co-ordinating research in the longer term. Specialists in water quality modelling from UK Water Authorities and Universities, the USA, Australia and Italy were invited. The programme of research was decided at an early stage and was divided into the following areas:

- methodological problems in modelling water quality;
- real-time forecasting of water quality in river systems;
- decision-making and control in water quality management;
- modelling and control of waste water treatment plants.

This report describes research carried out during the workshop.

REPORT NO 88

July 1984

CONTENTS

LIST OF PARTICIPANTS

Professor R C Spear
School of Public Health
Department of Biomedical and
Environmental Sciences
University of California
Berkeley
USA

Professor S Marsili-Libelli
Department of Electronic Engineering
University of Florence
ITALY

Professor George M Hornberger
Department of Environmental Sciences
University of Virginia
USA

Dr P Jowitt
Department of Civil Engineering
Imperial College
University of London
U.K.

Mr S Cook
Public Health Department
Imperial College
University of London
U.K.

Dr P E O'Connell
Institute of Hydrology
Wallingford
U.K.

Dr R Black
Department of Physics
Western Australia
Institute of Technology
Perth
AUSTRALIA

Dr M B Beck
IIASA
Laxanberg
Vienna
AUSTRIA

Mr R Williams
Institute of Hydrology
Wallingford
U.K.

KEY TO AUTHORS OF SECTIONS

Section	
1	P.G. Whitehead
2	P.G. Whitehead and M.B. Beck
3.1	P.G. Whitehead
3.2	R.J. Spear and G.M. Hornberger
3.3	S. Cook and P. Jowitt
4.1	P.G. Whitehead
4.2.1	R.Black
4.2.2	P.G. Whitehead and R. Williams
4.3	P.G. Whitehead and G.M. Hornberger
4.4	S. Marsili-Libelli and R.J. Spear
5.1	P.G. Whitehead
5.2	R.J. Spear and G.M. Hornberger
5.3	G.M. Hornberger and R.J. Spear
5.4	P. Jowitt
6	P.G. Whitehead
Appendix 1	P.G. Whitehead
Appendix 2	S. Cook

ACKNOWLEDGEMENTS

These proceedings testify to a considerable research effort by all of the participants; however, the Workshop could not have been held without the financial support of the Thames Water Authority. The personal interest and encouragement of D G Jamieson (Thames Water Authority) was much appreciated. The support of the Director of the Institute of Hydrology, Dr J S G McCulloch is also acknowledged.

Finally, thanks are due to Mrs J Patching for her efforts in typing these proceedings.

1. INTRODUCTION

In addition to being major sources of water, river systems are used as the principal disposal pathways for industrial and domestic effluent. As demand for water increases and the variety of pollutants becomes more diverse, there is a requirement for effective decision-making techniques which can be applied to water quality management problems. Accordingly over the past twenty years or so considerable effort has been devoted to the development of such techniques.

In his 1962 monograph on water pollution, the eminent water research economist, Allen Kneese, suggested that one of the major research needs in the area of water quality control was the development of "a methodology for keeping track of quality changes and quickly computing the concentration of pollutants (and significant associated variables such as dissolved oxygen) at all relevant points of use, as a function of a variety of conditioning factors (including) waste loads at particular outfalls, biological, chemical and physical conditions, and volume of stream flow" (kneese, 1962).

In order to meet these requirements several major studies were undertaken in the USA and in the UK. Initially planning models were developed, as in the case of the Delaware Study (1966) and the Trent Study (1968), and the models were used to investigate alternative capital investment programmes. Later studies, such as the Bedford Ouse Study (1979), considered the dynamic or short term variations in water quality and the models were applied to design or operational management problems (see Appendix 1). With the need to operate water and waste water treatment facilities efficiently the dynamic models have become increasingly useful in assessing alternative control measures. In addition, the sustained innovation of electronic engineering devices into the water/wastewater industry has markedly improved the capacity to implement good operational management practice. At the same time, longer-term changes in the nature of water pollution problems - for example, substantially more complex objectives for management, the increasing probability of accidental pollutant spillages and the changing role of treatment facilities - have increased the need for such good practice.

In Section 2 of these proceedings the role of systems analysis in water quality management is considered. It deals essentially with the new potential for operational management, although this is naturally linked with process design and, in turn, with the planning aspect of management. The policy implications of this changing management emphasis are important and the reader is also referred to the IIASA executive report on this subject (Beck, 1981).

In Section 3, two specific techniques used extensively in the workshop are described. The generalised sensitivity analysis approach developed by Spear and Hornberger is particularly relevant to 'ill-defined' environmental systems where the uncertainties associated with data, mechanisms and model parameters are often significant. The extended

Kalman filter is similarly a technique which accounts for the stochastic behaviour of a system and can be applied to a wide range of modelling problems.

In Section 4 several studies are described which illustrate the various stages of model development from nitrogen balance studies in river reaches through to detailed nitrogen budgets for river basins. The problems of modelling biological systems are emphasised in an algal modelling study and in an analysis of an activated sludge treatment plant.

In Section 5 aspects of control system design are investigated and, in particular, a new approach to control system design developed for environmental systems where there is considerable uncertainty associated with model structure and parameter estimates. One of the reasons for the limited application of control to environmental systems is that classical control techniques have often been applied without consideration of the underlying uncertainties. The techniques developed during the workshop, however, account for the stochastic nature of river systemsor treatment plants and represent a new development in this area.

Finally the appendices of the proceedings contain a computer listing and user notesfor the extended Kalman filter program developed during the workshop together with a description of the Bedford Ouse monitoring, modelling and forecasting system developed by the Institute of Hydrology.

2. SYSTEMS ANALYSIS FOR WATER QUALITY MANAGEMENT

2.1 Introduction

Our attitude towards mathematical models in the workshop has been that they are important functional components of management and are therefore most useful in a prescriptive sense to aid management decisions concerning design and operational control. There are several areas where mathematical models and systems analysis techniques are of value and these may be summarised as:

(a) economic analysis, in which fixed capital and variable operating costs are considered and the effects of transient crises, failures and uncertainties assessed;

(b) analysis of the interactions between the components of a water resource system, the reliability of the components and their sensitivity to accidents and failures;

(c) support services for operational management, in which monitored data on flow and quality are used to estimate models, provide forecasts of systems behaviour and aid operational management decisions;

(d) process control systems synthesis, involving the design and implementation of control systems.

In many cases, models are developed for "off line" studies of water quality management. In other words, the model is used as a basis for analysing the feasibility and type of operational management to be implemented in practice. In the study of "on-line" systems, data are received in real-time and operational management or control decisions are taken to solve an immediate problem. In this section of the proceedings the role of systems analysis techniques in both "off-line" and"on-line" problems is considered.

2.2 Economic analysis

The most common type of application of a model in water quality management has been in the determination of a minimum cost solution for the location, capacity, and (pollutant removal) efficiency of new (or expanded) facilities that would achieve specified river water quality standards (Trent Study, 1968, Delaware Study - see Spofford, 1976). Typically, facilities here would mean wastewater treatment plants, land disposal of effluent, low flow augmentation, and artificial instream aeration. Following the solution of such an initial problem, there have also been many applications of more detailed models for minimum cost design of, for example, sewer networks and wastewater treatment plants. These are clearly applications limited strictly to the planning and design phases of management, in which long-term fixed cost considerations have dominated exclusively.

By contrast, there have been relatively few studies on the economic trade-offs between capital costs and operating costs. As an example of this, over a decade ago it was suggested that seasonal waste treatment could result in substantial economic savings and that permission for variable waste control would allow trade offs between capital-intensive treatment facilities and facilities with high operation and maintenance costs. Neither suggestion appears to have been seriously considered, doubtless because energy costs prior to 1973 were relatively low and there was little incentive to reduce operating costs. With increasing energy costs however substantial savings are feasible. It is possible that seasonal waste treatment suggestions were unattractive because of theoretical and computational difficulties in applying optimization algorithms that would be able to handle the inevitable complexities of these issues. Since the earliest attempts at obtaining optimal solutions (investment cost minimization) to water quality management in the mid- to late-1960's, there has been extensive development of applicable methods of optimization. Thus we are now able to work within a framework for economic analysis where the planning of strategies for water quality management can address jointly both fixed and variable costs (Smeers, 1980). Furthermore, it is possible to recognize the time-variable and, to some extent, unpredictable character of meteorology and the receiving water body by including, for example, statistical distributions of the stream discharge. This kind of progress is most significant, for it greatly enlarges the scope of planning options, including the expected benefits of operational management, that can be assessed on an economic basis. Other factors which are then easily accommodated within such a broad framework include: uncertainty in the knowledge of stream behaviour; extreme or abnormal operational events, such as toxic substance spillages, or treatment plant bypassing and overflow, and coordination of pollutant removal facilities with low-flow augmentation facilities. The statistical, or probabilistic, nature of this economic analysis permits also a cost minimization subject to the satisfaction of probabilistic water quality standards such as, for example, the frequency of excessive nitrate-nitrogen concentrations persisting for a given period of time.

2.3 Interactions, reliability and sensitivity

The essence of a sensitivity analysis of any given configuration of components in a water quality system is the determination of the effects of and operational management response to natural fluctuations in water quality caused by varying hydrological effects and to sewage treatment failures and accidents such as pollutant spillages. It is in this sense of interaction among processes, that our understanding is weakest of all and where systems analysis techniques can be most useful. Mathematical models can be used to resolve the tradeoffs between the complex, multiple objectives that operational management has to address itself to. To what extent, for example, does the design of the system permit operational management to coordinate individual processes in order to confine the effects of a failure? For the time being we are probably

only able to consider a limited number of questions of this kind, with transient violations of nitrate-nitrogen standards being one example. In order to investigate the nitrate problem a multi-compartment model of the Thames River Basin has been developed and this is described in Section 4.2 of the proceedings.

The complexity of interactions between components in a water resource system may be illustrated by the following common uses of rivers (Jamieson, 1980):

(a) natural drainage : rivers are the means of disposing of excess rainfall;

(b) induced drainage : urbanisation and improved drainage of agricultural land has enhanced the natural drainage processes;

(c) water supply : rivers are used as the major sources of water for municipal, industrial and agricultural uses;

(d) navigation : minimum river depths are required to permit navigation for commercial and recreation uses;

(e) sewage disposal : rivers are used as the principal pathways for the disposal of industrial and domestic sewage and effluents;

(f) hydro-electric power : rivers and reservoirs are frequently used for power generation;

(g) fisheries : rivers provide the natural habitat for fish and the level of river pollution will determine the variety of species observed;

(h) amenity and recreation : water authorities are obliged under the 1973 Water Act to promote recreational and amenity activities;

(i) flood control : rivers need to be controlled to prevent flooding.

A river basin managed for one specific aspect may well provide benefits to other aspects but conflict with others. For example, disposal activities will conflict with water conservation and supply activities and minimum depths for navigation will conflict with the flood control requirements. Of course, different interests in river basin management will have different criteria for assessing performance and it is perhaps inappropriate to talk of 'optimal' operating rules for river basin management but rather to aim for a 'satisficing' approach (Jamieson, 1978). In this approach the state of the system is deemed satisfactory for specific interests provided it is within pre-defined boundaries for each activity. Systems analysis has a major role in assessing the trade-offs in such a situation and in evaluating the reliability of

each component such that the effects of failure can be ascertained.

At another level of detail, within each component of a water resource system there are interactions between physical, chemical and biological processes and these also require analysis using systems techniques. The identification of process models and the estimation of model parameters is an area of increasing study and during the workshop a range of time series analysis and filtering techniques have been employed. In particular, a new approach to generalised sensitivity analysis has been applied to the modelling of complex systems such as algal population growth dynamics in rivers and biological waste water treatment processes. The theoretical aspects of the systems techniques are described in section 3 of the proceedings, together with several applications in Section 4. The sensitivity analysis provides particularly useful information on the uncertainties associated with model parameters and hence process interactions. This information can be of direct benefit in subsequent management studies to translate model errors and the uncertainties on process interactions into model predictions.

2.4 Support services for operational management

Decisions taken in river basin management are seldom simple and up to date information is therefore required on the current state of the system together with forecasts of future system behaviour. In general, operational managers do not, at present, have the facilities to observe or forecast behaviour in real time. There are, however, exceptions to this such as in the case of the Bedford Ouse (see Appendix I) where an on-line flow and quality monitoring, telemetry and forecasting system has been established. Considerable benefits can accrue from such a system. For example, when evaluating the risks to river conditions associated with the loss of efficiency at an effluent treatment plant it is necessary to forecast the short term (ie day to day or hour to hour) changes in river water quality. A subjective approach can be taken whereby a pollution inspector draws on many years of local experience to assess a pollution event. An alternative approach is to use a computerised mathematical model to forecast flow and water quality. For example, it may be necessary to close a direct abstraction to a water supply treatment plant whilst a pulse of polluted water flows past. By having prior warning of the time of arrival and the severity of pollution conditions, it is possible to increase operational efficiency as well as safeguard water supply. Since a river is a complex hydrological system and 'dynamic' in character systems analysis techniques are required to develop realistic models and to use data in real time to forecast future behaviour.

As mentioned previously, there is a strong element of uncertainty associated with any estimate or forecast of water quality. It derives in part from an inadequate and imprecise knowledge of the system's behaviour and in part it is a function of the inevitable errors associated with measurements of water quality. The development of accurate and reliable sensors (providing less uncertain observations) for as many variables as possible is, therefore, an essential goal in the long term.

In our view, an operational monitoring system, which will inevitably have objectives different from those of a monitoring network for planning, should possess the following attributes:

(a) the provision of unambiguous and reliable measurements of short-term changes - because the primary concerns are those of diurnal variations, accidental events, and meteorological variability;

(b) the reliable measurement of "surrogate" variables supplemented by data processing algorithms - because such a combination fully exploits all the opportunities to convert reliable data into useful information. (As a typical example, knowledge of the state of "biological activity" is often desirable for the control of the activated sludge process; it could, in principle, be reconstructed through the combination of a model and processed data on the surrogate variables of substrate and metabolic end-product concentrations in the influent and effluent streams).

Thus systems techniques coupled to on-line monitoring and telemetry systems can provide a powerful new tool for operational management.

2.5 Process control systems synthesis

An ideal control system design project might have the following four distinct phases.

(a) design and implementation of experimental work and collection of experimental field data;

(b) derivation and verification of a mathematical model by reference to the field data;

(c) specification of process control objectives, and control system synthesis and evaluation by reference to the mathematical model;

(d) installation of the control system on the field unit.

The problems of control system design actually divide into two types: either a control scheme is to be designed for an existing facility, or the control scheme is to be developed together with the design of the facility prior to construction. In the chemical process and refining industries the latter approach is conventional. However this is generally not the case in the civil engineering field and it is often necessary to design controllers after construction. This increases the probability that such controller designs are only partially effective, which in turn may reinforce the view that operational control is not feasible. Given this situation, systems analysis has a major role to play in demonstrating the advantages of operational control.

Both the theory and application of control techniques have advanced during the past twenty years. In the sixties major advances were made in those industries, such as aerospace, nuclear power, chemical process, oil refining and paper where automation was required to ensure adequate

integrated control. The requirements of the four phases listed above for control system synthesis could also be relatively easily satisfied in these industries. Although conventional single loop three term controllers (ie controllers with proportional, integral and derivative action) were principally used at the local control level, sophisticated techniques were developed for the control of multivariable interacting systems and for the hierarchical control of systems. For example, in the refining industry the increased integration of component plants meant that in addition to localised control, supervisory or hierarchical control was required to optimise the overall plant performance. The theory of optimal control became a major area of research. In general, the theory has outpaced application and there are few optimal control systems that have been applied in practice. Rather, 'conventional' control techniques have been applied together with hierarchical control to provide an adaptive system with control targets for each sub-compartment being established from the central control.

In the water industry there has been little incentive to introduce control. Capital has always been available to build large treatment plants and the design of these has not been conducive to control techniques. Management has always had time to respond to any emergency because of the storage available and, prior to 1974, the fragmented nature of the water industry made control impossible. However with the 1974 reorganisation all the functions of the water industry listed in Section 2.3 are now the responsibility of each water authority. Integration has meant that control is now possible and with the recent developments in micro-computers there are financial advantages to be gained from increased automation.

Major trends are evident in present computer installations for process control: the previously dominant preference for a single, large-scale, central computer is being superseded by the emerging philosophy of dividing the computational burden among a host of small sub-units (each designed to carry out a specific set of tasks), whose basic component is the micro-processor. Microprocessors may be regarded as low-cost, flexible computing power that can be installed along a decentralized network. In such a context they should support a variety of activities: data acquisition and instrument management, data exchange with the central computer and communication-line management, peripheral process control and (control) actuator management. Over and above these "administrative and supervisory" functions, however, there are possibilities for applying microprocessors to various tasks of estimation, forecasting and control. The task might be as simple as the detection of and compensation for instrument drift, which would be extremely important for avoiding incorrect operating information and which extends our notion of a reliable monitoring network. Or the task might be as complex as the reconstruction of estimates of biological activity using a simple model of substrate/biomass interaction, and this task in turn could be embedded in a fully closed-loop process controller - where again the controller component could be programmed on a microprocessor. Given these possibilities, and indeed much research and development work remains to be done, one can discern certain strategic requirements for the development of models of water quality relationships. and the application of microelectronics to the water industry.

A recent proposal by Jamieson (1980) to develop an integrated hierarchical control system for the Thames Water Authority is an example of how important micro-electronics and the related techniques of systems analysis are to the future of the water industry.

2.6 Conclusions

As river basins become highly developed so the interactions and activities affecting and affected by water quality become more subtle and complex. The nature of the water quality problem often changes over time as the system alters. For example, problems perceived by management for most urgent attention change from easily degradable wastes, to point source pollution problems to non-point source pollution problems, to eutrophication problems, to toxic chemical problems and so on. Systems analysis techniques, therefore, have to cater for a wide range of problems and mathematical models must be chosen carefully to suit the nature of the particular water quality problem.

Past management strategies have lead to a progressively complex infrastructure of engineering facilities in the river basin. However, over the past decade there have been major innovations in providing cheap reliable electronic equipment which has been used to increase the monitoring and control capabilities of operational management. These improvements in technical hardware need to be matched by developments in systems analysis techniques. The overall objective of the workshop has been, therefore, to advance the theory of systems analysis and apply the techniques to significant problems and these are described in the following sections of the proceedings.

3. BASIC METHODOLOGY

3.1 Intoduction

In the analysis of water quality problems, mathematical modelling studies can sometimes aid in hypothesis development, in the screening of data and in the design of control schemes which can deal with the inherent uncertainty in such systems. Circumstances usually require models used in this way to be simulation models closely based on conceptual descriptions of component processes. As a result such models may contain many ill-defined parameters, a fact which severely limits the reliance that can be placed on the outcome of any single simulation. In an attempt to overcome this difficulty it has been proposed that parameters be assigned statistical distributions which reflect the degree of parametric uncertainty and that these distributions be used in Monte Carlo simulation analyses. In section 3.2, a variation on this theme is proposed in which the system's behaviour is stipulated initially and a classification algorithm defined and applied to the model output. This algorithm results in each simulation run being classified as a behaviour, B, or not a behaviour $\bar{B}$. The parameters leading to the result are stored according to the behavioural outcome. Subsequently, all parameter vectors are subjected to analysis to determine the degree to which the *a priori* distributions separate under the behavioural mapping. This separation, or lack thereof, forms the basis for a generalized sensitivity analysis in which parameters and their related processes important to the simulation of the behaviour are singled out.

In Section 3.3, the problem of estimating the parameters of a non-linear continuous/discrete water quality system is considered. The theoretical aspects of the extended Kalman filter technique are presented and a computer program developed during the workshop is listed in Appendix 2 together with computation and user notes.

3.2 A generalised sensitivity analysis technique

3.2.1 Introduction

Theories useful for developing mathematical models or for designing control systems are, for the most part, pertinent to well-defined systems, i.e., those for which a valid model structure is available and for which parameter values can be accurately specified. As Young (1978) has pointed out, strategies for building models of well-defined systems are rarely (or never) suitable for application to poorly defined systems in which uncertainties in measurements, model structure and parameter estimates are likely to exert a dominant influence.

Problems involving biological systems, such as denitrification processes, algal blooms in rivers or biological waste treatment processes, are often poorly-defined for a variety of reasons. Biological processes and complex chemical reactions that take place in these systems are not well understood, at least in quantitative terms. Data are often limited in quantity and quality and non-

stationarity is the rule rather than the exception. Nevertheless, the ultimate goal of many efforts relating to modelling such systems is to develop a firm basic understanding of processes and an ability to control these systems.

A technique, referred to as a generalized sensitivity analysis for the treatment of poorly defined systems, has been recently developed (Spear and Hornberger, 1980; Hornberger and Spear, 1981). The basis of the method is the utilization of a simulation model together with a classification algorithm. The classification allows any particular trajectory of the state variables of the system to be identified as either representative of the observed (or desired) behaviour of the system or not representative of the behaviour. The idea is to inject uncertainty into the simulation model of the system by specifying the parameters via probability density functions (rather than point estimates) and then to perform Monte-Carlo simulations choosing parameter values from the specified distributions. The result of each Monte-Carlo replication is classified as either a behavior, B, or as a non-behavior, $\bar{B}$. Subsequent to the Monte-Carlo trials, statistical analysis of the parameter vectors is used to isolate those parameters important in simulating the salient features defining the observed behaviour. The sensitivity rankings of the parameters are taken to be indicative of the relative importance of uncertainties in various component processes. The method has been applied to identify critical uncertainties in an estuarine eutrophication problem (Hornberger and Spear, 1980; Spear and Hornberger, 1980) and in an economic analysis of solar power satellites (Spear and Hornberger, 1981).

3.2.2 Model description

Assume a water quality or biological system to be modelled by a set of first order ordinary differential equations. Let these equations be given in the form:

$$\frac{d\underline{x}(t)}{dt} = \dot{\underline{x}}(t) = \underline{f}[\underline{x}(t), \underline{\xi}, \underline{z}(t)] \qquad (3.1)$$

where $\underline{x}(t)$ is the state vector and $\underline{z}(t)$ a set of time variable functions which include input or forcing functions. The vector $\underline{\xi}$ is a set of constant parameters described more fully below. Thus for $\underline{\xi}$, $\underline{z}(t)$ and $\underline{x}(o)$ specified, $\underline{x}(t)$ is the solution of the system of equations and is a deterministic or stochastic function of time as determined by the nature of $\underline{z}(t)$. For simplicity of exposition, $\underline{z}(t)$ will be treated hereafter as a deterministic function of t. Under this assumption, there are two types of uncertainty with which we will deal; uncertainty in the model structure, i.e. in the functions, $\underline{f}$, and uncertainty in the parameter values $\underline{\xi}$. Different model structures would pertain to competing hypotheses on system functioning (e.g., phosphorus limitation vs nitrogen limitation in a eutrophication problem); we use the term *scenario* to indicate a particular structure.

For a given scenario each element of the vector $\underline{\xi}$ is defined as a random variable, the distribution of which is a measure of our uncertainty in the 'real' but unknown value of the parameter. These parameter distributions are formed from data available from the literature and from experience with similar structures. For example, the literature suggests that the maximum growth rate of the algae

Chlorella vulgaris is almost certainly between 1.5 and 2.5 days^{-1} at water temperature near 25°C. Interpreting these limits as the range of a rectangularly distributed random variable, and forming similar *a priori* estimates for the other elements of $\underline{\xi}$, results in the definition of an ensemble of models for a given scenario. Some of these models will, we hope, mimic the real system with respect to the behaviour of interest.

3.2.3 The problem-defining behaviour

Turning now to the question of behaviour, recall that for a given scenario every sample value of $\underline{\xi}$, drawn from the *a priori* distribution, results in a unique state trajectory, $\underline{x}(t)$. Following the usual practice, we assume that there is a set of observed variables $\underline{y}(t)$, which can be calculated from the state vector important to the problem at hand. So, for each randomly chosen parameter set ξ^*, there corresponds a unique observation vector $\underline{y}^*(t)$. Since the elements of $\underline{y}(t)$ are observed (ie they are measured in the real system) it is sensible to define behaviour in terms of $\underline{y}(t)$. For example, suppose y_i is the concentration of phytoplankton in a body of water and the problem in question concerns unwanted algal blooms due to nutrient enrichment. Then there is some value of y_i above which a bloom is defined to have occurred and the behaviour is defined by this critical value.

In general a number of behaviourcategories can be used. Without loss of generality, however, we can consider the case for which behaviour is defined in a binary sense, that is, it either occurs or does not occur for a given scenario and set of parameters $\underline{\xi}$. It follows that a rule must be specified for determining the occurrence or non-occurence of the behaviour on the basis of the pattern $\underline{y}(t)$. It is also possible that the behaviour might depend on the vector $\underline{z}(t)$. For example, suppose one element of $\underline{z}(t)$ were water temperature. We might be interested only in extreme values of $\underline{y}(t)$ when adjusted or controlled for temperature variations. In any event, the detailed definition of behavior is problem-dependent and, for present purposes, it is sufficient to keep in mind that a set of numerical values of $\underline{\xi}$ leads to a unique time function $\underline{y}(t)$ which, in turn, determines the occurence or non-occurrence of the behaviour conditions, perhaps, by $\underline{z}(t)$.

3.2.4. Application of the technique

We have now presented the class of models to be studied, defined the scenario concept and described how we propose to deal with parametric uncertainty. For a given scenario behaviour and set of parameter distributions $\underline{\xi}$, it is possible to explore the properties of the ensemble via computer simulation studies. In particular, a random choice of the parameter vector $\underline{\xi}$ from the predefined distributions leads to a state trajectory $\underline{x}(t)$, an observation vector $\underline{y}(t)$ and, via the behaviour defining algorithm to a determination of the occurrence or non-occurrence of the behaviour. A repetition of this process for many sets of randomly chosen parameters results in a set of sample parameter vectors for which the behaviour was observed and a set for which the behaviour was not observed. The key idea is then to attempt to identify the subset of physically, chemically or biologically meaningful parameters which appear to account for the occurrence or non-occurrence of the behaviour . More traditional

sensitivity analyses of large ecological models inevitably show that a surprisingly large fraction of the total number of parameters is simply unimportant to the critical model behavior. We maintain that this unimportant subset or conversely the critical subset, may be tentatively specified rather early in any study.

Ranking the elements of $\underline{\xi}$ in order of importance in the behavioral context is accomplished through an analysis of the Monte-Carlo results. The essential concept can be best illustrated by considering a single element, ξ_k, of the vector $\underline{\xi}$ and its *a priori* cumulative distribution as shown in Figure 3.1. Recall that the procedure is to draw a random sample from this parent distribution (a similar procedure is followed for all other elements of $\underline{\xi}$), run the simulation with this value and record the observed behaviour and the total vector $\underline{\xi}$ therewith associated. A repetition of this procedure results in two sets of values of ξ_k, one associated with the occurrence of the behaviour B, and the other with non-behaviour, $\bar{B}$. That is, we have split the distribution $F(\xi_k)$ into two parts as indicated in the figure. This particular example would suggest that ξ_k was important to the behavior since $F(\xi_k)$ is clearly divided by the behavioural classification. Alternatively, if the simple values under B and $\bar{B}$ appeared both to be from the original distribution $F(\xi_k)$ then we would conclude that ξ_k was not important.

3.2.5. Sensitivity ranking of parameters

For the case where $\underline{z}(t)$ is a deterministic function of time, the parameter space is cleanly divided by the behavioural mapping; that is, there is no ambiguity regarding whether a given parameter vector results in B or $\bar{B}$. Our analysis then focuses on the determination of which parameters or combinations of parameters are most important in distinguishing between B and $\bar{B}$. We will restrict the discussion to the case for which the parameter vector mean is zero and the parameter covariance matrix is the identity matrix. (A suitable transformation can always be found to convert the general problem to this case.) The problem of identifying how the behavioural mapping separates the parent parameter space can then be approached by examining induced mean shifts and induced covariance structure.

For example, we can base a sensitivity ranking on a direct measure of the separation of the cumulative distribution functions, $F(\xi_k|B)$ and $F(\xi_k|\bar{B})$. In particular, we utilize the statistic

$$d_{m,n} = \sup_x \left| S_n(x) - S_m(x) \right|$$

where S_n and S_m are the sample distribution functions corresponding to $F(\xi_k|B)$ and $F(\xi_k|\bar{B})$ for n behaviours and m non-behaviour. The statistic $d_{m,n}$ is that used in the Kolmogorov-Smirnov two sample test and both its asymptotic and small sample distributions are known for any continuous cumulative distribution function $F(\xi_k|B)$ and $F(\xi_k|\bar{B})$. Since S_n and S_m are estimates of $F(\xi_k|B)$ and $F(\xi_k|\bar{B})$, we see that $d_{m,n}$ is the maximum vertical distance between these two curves and the statistic is, therefore, sensitive not only to differences in central tendency but to any difference in the distribution functions. Thus, large values of $d_{m,n}$ indicate that the paramenter is important for simulating the behaviour and, at least in cases where induced covariance

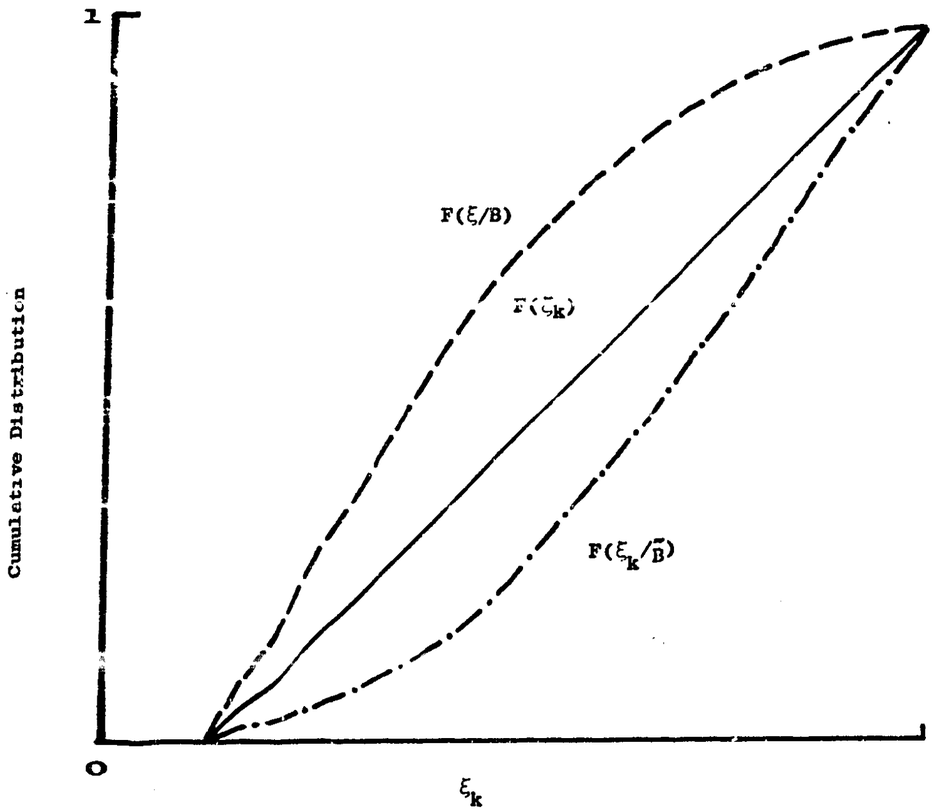


Figure 3.1. Cumulative distribution functions for parameter ξ_k $F(\xi_k)$ = parent, *a priori* distribution, $F(\xi_k/B)$ = distribution of ξ_k in the behaviour category. $F(\xi_k/\bar{B})$ = distribution of ξ_k in the non-behaviour category.

is small, the converse is true for small values of that statistic.

In general, however, ranking on the basis of the separation in the distribution functions along the original axes of the parameter space (the individual parameter values) is not sufficient. It is possible, for example, that the first and second moments for a single parameter might exhibit no separation and yet this parameter could be crucial to a successful simulation by virtue of a strong correlation with other parameters under the behaviour. For example, Figure 3.2 depicts a two-dimensional parameter space for which the cumulative distributions would not separate under the behavioural classification. Nevertheless, both parameters are important in determining whether the behaviour occurs. Clearly, it is the interaction between parameters which is crucial and information on the covariance between the two parameters will give insight into the degree of sensitivity in a case such as this. In fact, as shown more formally below, inspection of the covariance matrices of the parameter vectors in the two classes can provide important clues in assessing sensitivity.

This notion can be formalized as follows. Let $\underline{\xi}$ be the parameter vector. Since these vectors were normalized to have zero mean, unit variance and zero covariance it follows that

$$E(\underline{\xi}\,\underline{\xi}^T) = I = P(B)E(\underline{Z}_1\underline{Z}_1^T) + P(B)\,\underline{\mu}_1\underline{\mu}_1^T$$

$$+ P(\bar{B})E(\underline{Z}_2\underline{Z}_2^T) + P(\bar{B})\,\underline{\mu}_2\underline{\mu}_2^T \qquad (3.2)$$

where $P(B)$ and $P(\bar{B})$ are the probabilities of obtaining the behaviour and of not doing so, respectively;

ξ_1 is a parameter vector associated with B;
ξ_2 is a parameter vector associated with $\bar{B}$;
$\underline{\mu}_1 = E(\underline{\xi}_1)$
$\underline{\mu}_2 = E(\underline{\xi}_2)$
$\underline{Z}_1 = \underline{\xi}_1 - \underline{\mu}_1$
$\underline{Z}_2 = \underline{\xi}_2 - \underline{\mu}_2$
E is the expectation operator.

The case illustrated in Figure 3.2 suggests that incidences in which separation is not indicated in the univariate analyses should be singled out in the multivariate procedure. Assume that m of the distributions $F(\xi_k)$ did not separate under the behavioral mapping. Then $\mu_{1k} = \mu_{2k} = 0$ for each of these distributions. For two parameters for which no mean shift is observed (say ξ_i and ξ_j) the ij^{th} elements of $\mu_1\mu_1^T$ and $\mu_2\mu_2^T$ are zero and, according to equation 3.2, the corresponding off diagonal elements of the covariance matrices are such that

$$P(B)E(Z_{1i}Z_{1j}) = -P(\bar{B})E(Z_{2i}Z_{2j})$$

where $i \neq j$. Therefore, if a distribution does not separate under the behavioural mapping but does show induced covariance, the situation depicted in Figure 3.2, this covariance will be seen in both the covariance matrices under B and $\bar{B}$ and the magnitudes of the covariances

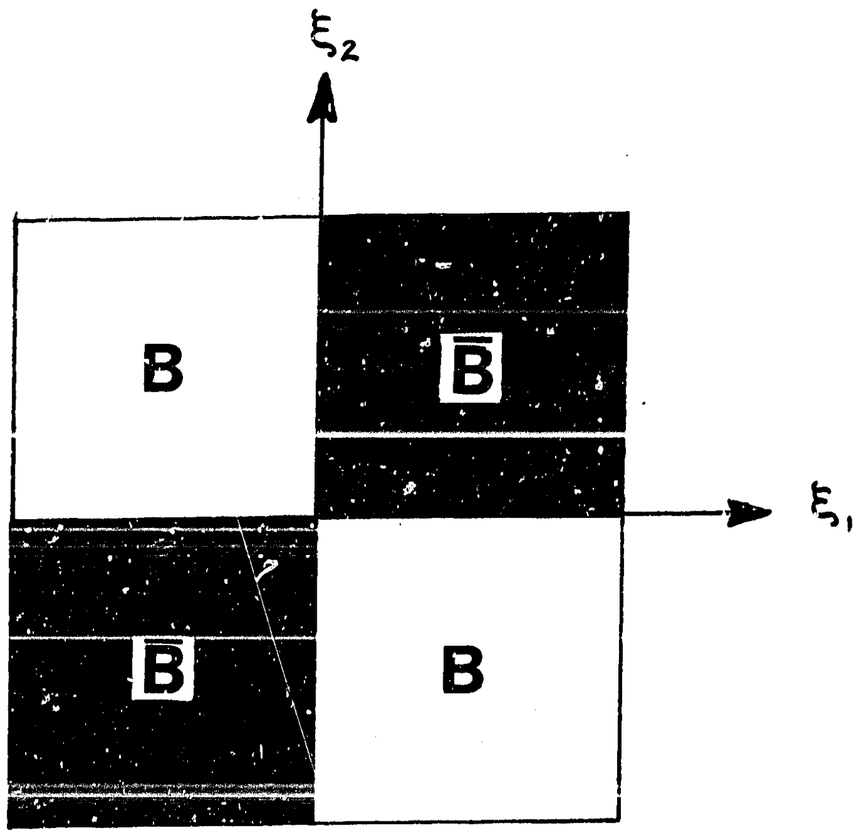


Figure 3.2. Schematic diagram of a two parameter case for which separation under the behavioural classification is total but for which discrimination by univariate tests is not possible.

will be related as indicated above. This is easily checked by inspection of the two matrices.

The problem of identifying "important" parameters in a situation where induced covariance is significant (e.g., the case shown schematically in Figure 3.2) can also be addressed by considering an analogy with the problem of discriminant analysis in the face of negligible differences in means between groups. Kendall and Stuart (1969) suggest that a principal components transformation may be useful for such problems. In fact, for the case depicted in Figure 3.2 it is intuitively obvious that the principal components of the covariance matrix of parameters in the behaviour class (or of those in the non-behavior class) define a new set of co-ordinate axes at 45° from the original ξ_1, ξ_2 axes. The cumulative distributions of the projections of the parameter vectors on these principal axes would indeed separate and the value of the $d_{m,n}$ statistic for these would again provide a useful measure of the separation.

In a more general case the behavioural classification would result in both separation along some of the original axes and in induced covariance. A principal-components transformation of the covariance matrix of either the behaviour class or the non-behavior class *defined with respect to the grand mean* can be used to advantage when information on both mean and covariance differences is important (Fukunaga and Koontz, 1979). That is, equation 3.2 can be written

$$E(\underline{\xi}\underline{\xi}^T) = I = P(B)E(\underline{\xi}_1\underline{\xi}_1^T) + P(\bar{B})W(\underline{\xi}_2\underline{\xi}_2^T) \qquad (3.3)$$

and if T is the matrix that diagonalizes the covariance matrix under the behaviour, $E(\underline{\xi}_1\underline{\xi}_1^T)$, then the same transformation must diagonalize the covariance matrix under the non-behaviour classification, $E(\underline{\xi}_2\underline{\xi}_2^T)$:

$$T^T E(\underline{\xi}\underline{\xi}^T)T = I = P(B)T^T E(\underline{\xi}_1\underline{\xi}_1^T)T + P(\bar{B})T^T E(\underline{\xi}_2\underline{\xi}_2^T)T,$$

or

$$I = P(B)\Lambda_1 + P(\bar{B})\Lambda_2 \qquad (3.4)$$

where Λ_1 and Λ_2 are diagonal matrices with the eigenvalues of the respective covariance matrices as the diagonal elements. The columns of the matrix T are eigenvectors of covariance matrices and the (normalized) components of those vectors are the direction cosines of the transformed axes relative to the original parameter axes. Thus, the projections of the parameters onto a transformed axis exhibit significant separation under the classification in terms of the $d_{m,n}$ statistic; the weights on individual parameters in that eigenvector indicate the importance of each parameter in explaining the separation.

Details of the generalized sensitivity analysis described above are perhaps best appreciated in the context of an actual problem in water quality modelling and/or control. Several examples are presented later in the proceedings and should serve to clarify the general approach.

3.3 Estimation in non-linear continuous/discrete water quality systems

3.3.1 Introduction

The problems of water quality modelling as emphasised in the previous section can often be addressed by a dynamic, state-space formulation; this leads to an internally descriptive model in which measurements are functions of the state space and not the input. The measurement functions will then be of the form:

$$dy(t) = h(x(t),t)dt.$$

Errors in the observations may be incorporated by the addition of a measurement noise term,

$$dy(t) = h(x(t),t)dt + d\eta(t)dt$$

Lack of knowledge of the process dynamics and random influences may be incorporated in the system equation by the addition of a system noise term,

$$dx(t) = f(x(t),u(t),t)dt + d\xi(t)$$

or by,

$$dx(t) = f(x(t),u(t),t,\xi(t))dt$$

where

x(t) is the n vector of state variables;

u(t) is the ℓ vector of deterministic inputs;

y(t) is the m vector of observations.

The state of the system and the parameters are normally unknown since we observe the system through a set of noisy measurements. The problem to be addressed is therefore one of state and parameter estimation through the examination of system responses to a set of deterministic (or comparatively noise-free) inputs.

3.3.2 Recursive estimation and the extended Kalman filter (EKF) algorithms

For on-line control and real-time forecasting situations, data handling and storage requirements are minimised by using a recursive approach to the estimation problem. The recursive algorithms to be considered here permit the identification of time-varying parameters; this is useful in

model structure identification and may highlight previously un-modelled parameter/state or parameter/input relationships (Whitehead, 1979). The models employed for natural systems are often non-linear either in the system or measurement equations or both. Thus the estimator requirements are those of recursive state and parameter estimation in dynamic, non-linear, noise-corrupted, state-space systems from a set of noisy observations.

Many water quality systems are best described by continuous time equations; however the output observations are often sampled discretely and a continuous/discrete time approach is then required. The Kalman filter provides an optimal solution to the linear estimation problem (see Jazwinski, 1970 for a theoretical development of the Kalman filter). In the area of water quality modelling mechanistic processes are generally non-linear and an extended version of the Kalman filter is required.

Consider the continuous/discrete time system,

$$\begin{aligned} \frac{d\underline{x}(t)}{dt} &= f(\underline{x}(t), \underline{u}(t), t) + \underline{\xi}(t) \\ \underline{y}(t_k) &= h(\underline{x}(t_k), \underline{u}(t_k), t_k) + \underline{\eta}(t_k) \end{aligned} \qquad (3.5)$$

where f and h may be non-linear functions. The system equations may be linearised so that the Kalman filter algorithm can be used in the estimation. The linearisation is carried out via a 1st order Taylor series expansion of the system equations. This leads to the following linear differential equation (see Beck, 1979),

$$\begin{aligned} \delta\dot{\underline{x}}(t) &= F'(\bar{\underline{x}}(t), \bar{\underline{u}}(t), \delta t)\, \underline{x}(t) + G'(\bar{\underline{x}}(t), \delta\bar{\underline{u}}(t), t)\, \underline{u}(t) + \underline{\xi}(t) \\ \delta\underline{y}(t_k) &= H'(\bar{\underline{x}}(t_k), \bar{\underline{u}}(t_k), t_k)\delta\underline{x}(t_k) + \underline{\eta}(t_k) \end{aligned} \qquad (3.6)$$

which is in the required form for the Kalman filter and where the elements of the Jacobian matrices F', G', H' are defined as follows:

$$F'(\bar{\underline{x}}(t),\bar{\underline{u}}(t),t)_{i,j} = \frac{\partial f_i(\bar{\underline{x}}(t),\bar{\underline{u}}(t),t)}{\partial x_j(t)}$$

$$G'(\bar{\underline{x}}(t),\bar{\underline{u}}(t),t)_{i,j} = \frac{\partial f_i(\bar{\underline{x}}(t),\bar{\underline{u}}(t),t)}{\partial u_j(t)}$$

$$H'(\bar{\underline{x}}(t_k),\bar{\underline{u}}(t_k),t_k)_{i,j} = \frac{\partial h_i(\bar{\underline{x}}(t_k),\bar{\underline{u}}(t_k),t_k)}{\partial x_j(t_k)}$$

and

$$\delta\underline{x}(t) \triangleq \underline{x}(t) - \bar{\underline{x}}(t)$$

$$\delta\underline{u}(t) \triangleq \underline{u}(t) - \bar{\underline{u}}(t)$$

The quantities $\bar{\underline{x}}(t)$ and $\bar{\underline{u}}(t)$ form a reference trajectory, notably

$$\dot{\bar{\underline{x}}}(t) = f(\bar{\underline{x}}, \bar{\underline{u}}, t).$$

The linearisation gives rise to a form of the Kalman filter which can be extended to cover non-linear problems. The resulting extended Kalman filter (EKF) algorithm consists of prediction equations and correction equations as follows:

$$\hat{\underline{x}}(t_{k+1}|t_k) = \hat{\underline{x}}(t_k|t_k) + \int_{t_k}^{t_{k+1}} f(\hat{\underline{x}}(t|t_k), \underline{u}(\tau), \tau)d\tau \qquad (3.7)$$

$$P(t_{k+1}|t_k) = \phi(t_k, t_{k+1}, \hat{\underline{x}}(t_k|t_k))P(t_k|t_k)\phi^T(t_k, t_{k+1}, \hat{\underline{x}}(t_k|t_k)) + \int_{t_k}^{t_{k+1}} \phi(t_k, \tau, \hat{\underline{x}}(t_k|t_k))Q_o\phi^T(t_k, \tau, \hat{\underline{x}}(t_k|t_k))d\tau \qquad (3.8)$$

where the transition matrix is

$$\phi(t_k, t_{k+1}, \hat{\underline{x}}(t_k|t_k)) = \exp\{\int_{t_k}^{t_{k+1}} F'(\bar{\underline{x}}(\tau|t_k), \underline{u}(\tau), \tau)d\tau\}$$

and where Q is the variance-covariance matrix of the system noise, $\underline{\xi}(t)$

Upon receipt of the observation vector $\underline{y}(t_{k+1})$, the following correction equations apply for the Kalman gain, the state and the covariance

$$K(t_{k+1}) = P(t_{k+1}|t_k)H'^T(t_{k+1})[(H'(t_{k+1})P(t_{k+1}|t_k)H'^T(t_{k+1}) + R(t_{k+1})]^{-1} \qquad (3.9)$$

$$\hat{\underline{x}}(t_{k+1}|t_{k+1}) = \hat{\underline{x}}(t_{k+1}|t_k) + K(t_{k+1})(\underline{y}(t_{k+1}) - h(\hat{\underline{x}}(t_{k+1}|t_k), \underline{u}(t_k), t_{k+1})) \qquad (3.10)$$

$$P(t_{k+1}|t_{k+1}) = [I - K(t_{k+1})H'(t_{k+1})]P(t_{k+1}|t_k) \qquad (3.11)$$

where R is the variance-covariance matrix of the measurement noise, $\eta(t)$. Independence between $\underline{\eta}(t_k)$ and $\underline{\xi}(t_k)$ is assumed. Further assumptions implicit in the algorithm are

(i) the reference trajectory is reset to the best state estimate at each sampling instant, this being $\hat{\underline{x}}(t_k|t_k)$.

(ii) the a priori state estimates are unbiased so that the a posteriori state estimates are also unbiased (Beck, 1979)

(iii) the reference input $\bar{\underline{u}}(t)$ is taken as equivalent to $\underline{u}(t)$.

3.3.3 State/parameter estimation

In most water quality modelling problems it is necessary to estimate both model states and model parameters. Process mechanisms are generally 'ill-defined' and parameters are unlikely to be known precisely. Whilst a priori estimates of parameters may be available from laboratory experiments or theoretical considerations it is necessary to derive field data using the estimation algorithms.

For the purpose of parameter estimation the system dynamics are written as

$$\dot{\underline{x}}(t) = f(\underline{x}(t), \underline{u}(t), t, \underline{\alpha}(t)) \qquad (3.12)$$

where

$$\underline{x}(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T$$

$$\underline{\alpha}(t) = [\alpha_1(t), \alpha_2(t), \ldots, \alpha_p(t)]^T$$

so that $\underline{\alpha}(t)$ is the vector of parameters to be estimated. We may combine the state and parameter vectors in the following way,

$$\underline{x}^*(t) = [x_1(t), x_2(t), \ldots, x_n(t), \alpha_1(t), \alpha_2(t), \ldots, \alpha_p(t)]^T \qquad (3.13)$$

where $\underline{x}^*(t)$ is known as the augmented state vector and may be employed in the system dynamics in place of $\underline{x}(t)$ and $\underline{\alpha}(t)$. In applying the EKF an assumption is made on the parameter dynamics and commonly employed parameter dynamics are,

(i) non-dynamic,

$$\dot{\underline{\alpha}}(t) = 0 ;$$

(ii) random walk,

$$\underline{\alpha}(t_{k+1}) = \underline{\alpha}(t_k) + \underline{\zeta}(t_k)$$

where $\zeta(t_k)$ is a Gaussian white noise process ;

(iii) Gauss-Markov,

$$\underline{\alpha}(t_{k+1}) = \beta\underline{\alpha}(t_k) + \underline{\zeta}(t_k)$$

$$-1 < \beta < 1$$

The EKF algorithm can now be employed for combined state and parameter estimation by substituting the augmented state vector $\underline{x}^*(t)$ for $\underline{x}(t)$ in equations (3.7)-(3.11). The terms state and parameter are semantic devices in this context since in estimation terms they are treated identically by the algorithm. This may create system non-linearities in an otherwise linear system. For example given

$$\dot{x}(t) = \alpha x(t)$$

then if,

$$x_1 \triangleq x$$

$$x_2 \triangleq \alpha$$

we have,

$$\dot{x}_1(t) = x_2(t)\, x_1(t)$$

$$\dot{x}_2(t) = 0$$

and the system equation is non-linear in $\underline{x}(t)$.

A useful means of avoiding induced non-linearities is to consider the states as deterministic and estimate the parameters only. In the case of the above example

$$x_1 \triangleq \alpha$$

$$\dot{x}_1 = 0$$

$$H(t) \triangleq x(t)$$

then,

$$y(t) = H(t)x_1(t) + \eta(t)$$

The measurement matrix is now time-varying. This linear systems formulation lends itself to Kalman filtering.

Lack of optimality in the EKF will occur

(i) when the functions h and f are not well approximated by a 1st order Taylor series expansion and/or the magnitudes of $\underline{\delta x}^{*2}$ (and higher order terms) are large. The Gaussian 2nd order filter attempts to overcome this problem (see Jazwinski, 1970).

(ii) when the initial augmented state vector is biased. This leads to inconsistent estimates and the filtered states may diverge from their true values. It is often useful to consider a limitation of the memory of the filter so that the learning process does not proceed at an unrealistic rate. This may conveniently be accomplished by using an exponentially age-weighted filter (Jazwinski 1970).

(iii) when the Q and R matrices are specified incorrectly the estimates will be biased and the EKF will be sub-optimal (Ljung, 1979). This is the major drawback of the EKF technique since accurate estimates of Q and R are generally not available.

In the solution of the EKF there are a number of computational problems to be overcome and these are discussed in detail in Appendix II together with a listing of the EKF PROGRAM AND ASSOCIATED USER NOTES.

The EKF has been applied extensively during the workshop (see Section 4.3) and in order to illustrate the technique an application to the modelling of an activated sludge treatment plant is considered here.

3.3.4 Application of the EKF to a waste water treatment model

In order to demonstrate the use and performance of the EKF in combined state and parameter estimation, a dynamic waste water treatment model is employed in simulation studies. A detailed description can be found in (Marsili-Libelli, 1980a, b); a short description will be useful here.

The model describes the behaviour in time of dissolved oxygen in a conventional activated sludge plant. The systems representation is shown in Figure 3.3.

where,

Tank 1	=	aeration tank
Tank 2	=	clarifier
u	=	artificial air input (Nm^3/hr)
Q	=	input flow rate (m^3/hr)
r	=	recycle ratio
V	=	aeration tank volume (m^3)
w	=	wastage ratio
S	=	pollutant concentration in tank 1 (mg/l)
C	=	biomass concentration in tank 1 (mg/l)
O	=	dissolved oxygen concentration in tank 1 (mg/l)
S_i	=	influent pollutant concentration (mg/l)
C_r	=	recycle biomass concentration (mg/l)

The aeration tank is modelled as a CSTR (continuously-stirred tank reactor) and the clarifier is assumed to be non-dynamic. A mass balance across the system gives the following set of differential equations:

$$\frac{dO(t)}{dt} = K_L (O_s - O(t))u(t) - \alpha k_2 S(t)C(t) - \delta k_3 \frac{C^2(t)}{S(t)} - \beta(t)(1+r)O(t) + \beta(t)O_i(t)$$

$$\frac{dC(t)}{dt} = k_2 S(t) C(t) - k_3 \frac{C^2(t)}{S(t)} - \beta(t)(1+r)C(t) + \beta(t)rC_r(t)$$

$$\frac{dS(t)}{dt} = -k_1 S(t)C(t) - \beta(t)(1+r)S(t) + \beta(t)S_i(t)$$

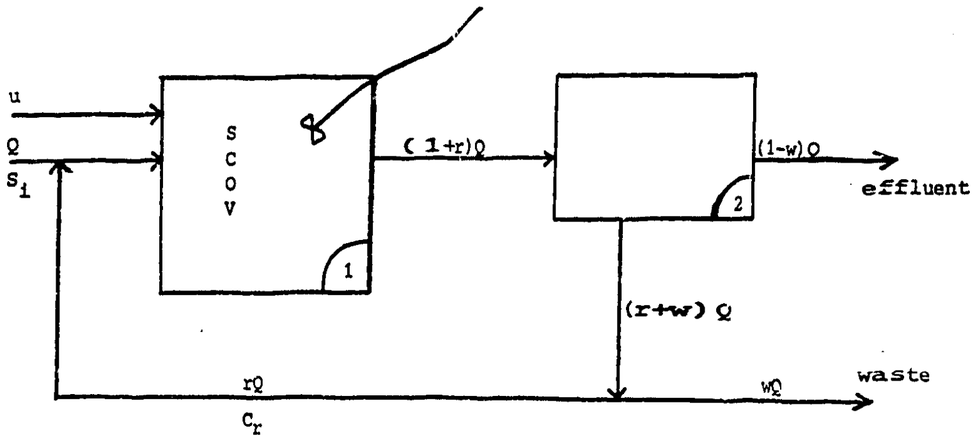


Figure 3.3 Schematic Diagram of Aeration Tank and Clarifier

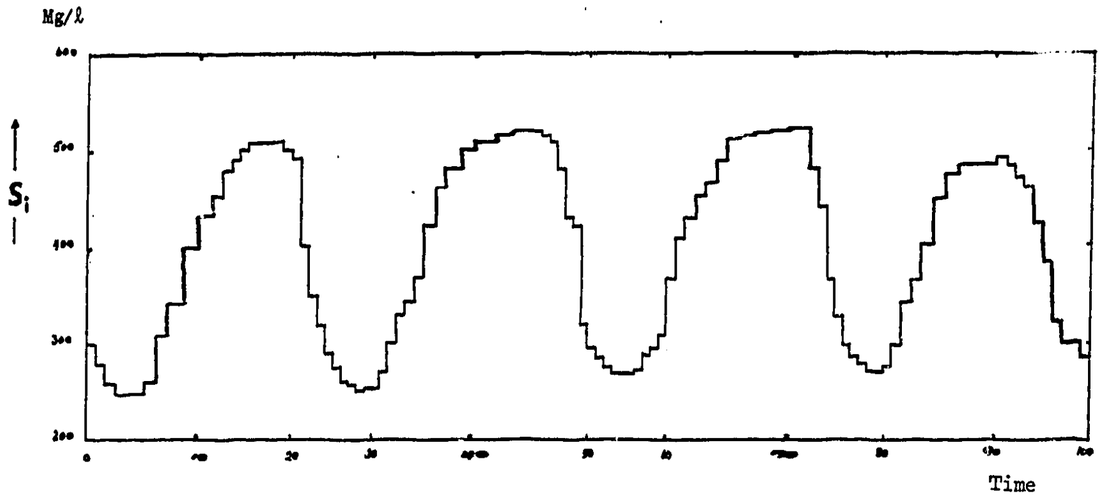


Figure 3.4 Input Pollutant Concentration

where,

K_L = artificial aeration transfer coefficient

O_s = dissolved oxygen saturation level

$O_i(t)$ = dissolved oxygen of incoming sewage

α, δ = yield coefficients

k_1, k_2 = synthesis rates ($mg^{-1}l\ hr^{-1}$)

k_3 = decay rate ($mg^{-1}l\ hr^{-1}$)

$\beta(t) = Q(t)/V$

Although this model has a number of simplifying assumptions, it has been shown (Marsili-Libelli, 1980a) to simulate plant behaviour reasonably well and, perhaps more important, it is of sufficient order and non-linearity to enable the capability of the EKF in the real-time forecasting of water quality to be assessed.

Simulation

The following assumptions are made in the simulation:

(1) $O_i(t) = 1\ mg\ l^{-1}$

(2) $\beta(t) = 0.14\ h^{-1}$

(3) $C_r(t) = 10^4\ mg\ l^{-1}$

(4) $O_s = 8.65\ mgl^{-1}$

(5) $r = 0.5$

(6) $u(t) = 12000\ N\ m^3\ hr^{-1}$

The system is forced by the input pollutant concentration $S_i(t)$ as shown in Figure 3.4. The pollutant can be considered as a measure of carbonaceous material (eg BOD). Unmodelled uncertainty in the system equations is introduced as an additive random noise term, i.e.

if $d\underline{x}(t) = f(\underline{x}(t), u(t), t)\, dt + d\beta$

$\underline{x} = [x_1, x_2, \ldots, x_7]^T$

$x_1(t) \triangleq O(t)$ $\qquad$ $x_4(t) \triangleq k_1$

$x_2(t) \triangleq C(t)$ $\qquad$ $x_5(t) \triangleq k_2$

$x_3(t) = S(t)$ $\qquad$ $x_6(t) \triangleq k_3$

$x_7(t) \triangleq k_L$

$\underline{d\beta}$ = vector Brownian motion process

then,

$$\underline{x}(t+1) = \underline{x}(t) + \int_{t}^{t+1} f(\underline{x}(\tau),u(\tau),\tau)d\tau + \underline{\xi}(t)$$

$\underline{\xi}(t)$ = vector of realisations from a zero-mean, normal distribution such that,

$$E[\xi_i(t)\xi_j(t)] = \sigma^2_{\xi i}\,\delta_{ij}$$

with

$$\delta_{ij} = 1 \quad i = j$$
$$= 0 \quad i \neq j$$

Thus, $\underline{x}(t+1)$ gives the vector of 'true' state variables.

Observations:

In current waste water treatment practice, dissolved oxygen is monitored continuously. Similarly, MLSS (mixed liquor suspended solids) which is often regarded as being proportional to the biomass concentration, can be measured relatively quickly. This is not so for the pollutant concentration (although improved techniques are becoming available). In the real-time forecasting sense, we may have available only observations of O(t) and C(t), but not S(t).

Since the system is forced by deterministic observations of influent pollutant, the model cannot be used in real-time forecasting. This argument, whilst presently true, is unimportant since the choice of observed variables is intended to illustrate a situation where combined state estimation and reconstruction is necessary.

The observations are created through

$$\underline{y}(t) = H\,\underline{x}(t) + \underline{\eta}(t)$$

where

$$\underline{y}(t) = [y_1(t),\ y_2(t)]^T$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$\underline{\eta}(t)$ = vector of realisations from a zero-mean, normal distribution such that,

$$E[\eta_i(t)\,\eta_j(t)] = \sigma^2_{\eta_i}\,\delta_{ij}$$

A hundred data points were synthesised using

$$Q = \begin{vmatrix} 0.0095 & & & & & & \\ & 20.9764 & & & & & \\ & & 1.1664 & & & & \\ & & & 0 & & & \\ & & & & 0 & & \\ & & & & & 0 & \\ & & & & & & 0 \end{vmatrix} \quad \text{and} \quad R = \begin{vmatrix} 0.1115 & \\ & 506.2500 \end{vmatrix}$$

Estimation

The parameters k_1, k_2, k_3 and K_I are estimated as non-dynamic states X_4, X_5, X_6 and X_7. The pollutant concentration S(t) is reconstructed as state X_3 from filtered estimates of O(t) and C(t) defined as states X_1 and X_3.

The one-step-ahead forecasts of dissolved oxygen and pollutant obtained from the continuous/discrete EKF are shown together with the synthesised values in Figure 3.5. The initial state error covariance matrix was

$$P_o = \begin{vmatrix} 2 & & & & & & \\ & 100 & & & & & \\ & & 5 & & & & \\ & & & 3 & & & \\ & & & & 0.4 & & \\ & & & & & 0.6 & \\ & & & & & & 2 \end{vmatrix}$$

and the elements of the matrices Q and R were taken as given above. The initial parameter estimates were pertubed by 50% of their true values and this leads to bias in the forecasts up to sample 25 when parameter convergence is achieved. The parameter estimates together with their true values are shown in Figure 3.6. A further run was performed with P_o and R as given above but with

$$Q = \begin{vmatrix} 0.0095 & & & & & & \\ & 20.9764 & & & & & \\ & & 1.1664 & & & & \\ & & & 0.05 & & & \\ & & & & 0.005 & & \\ & & & & & 0.05 & \\ & & & & & & 0.005 \end{vmatrix}$$

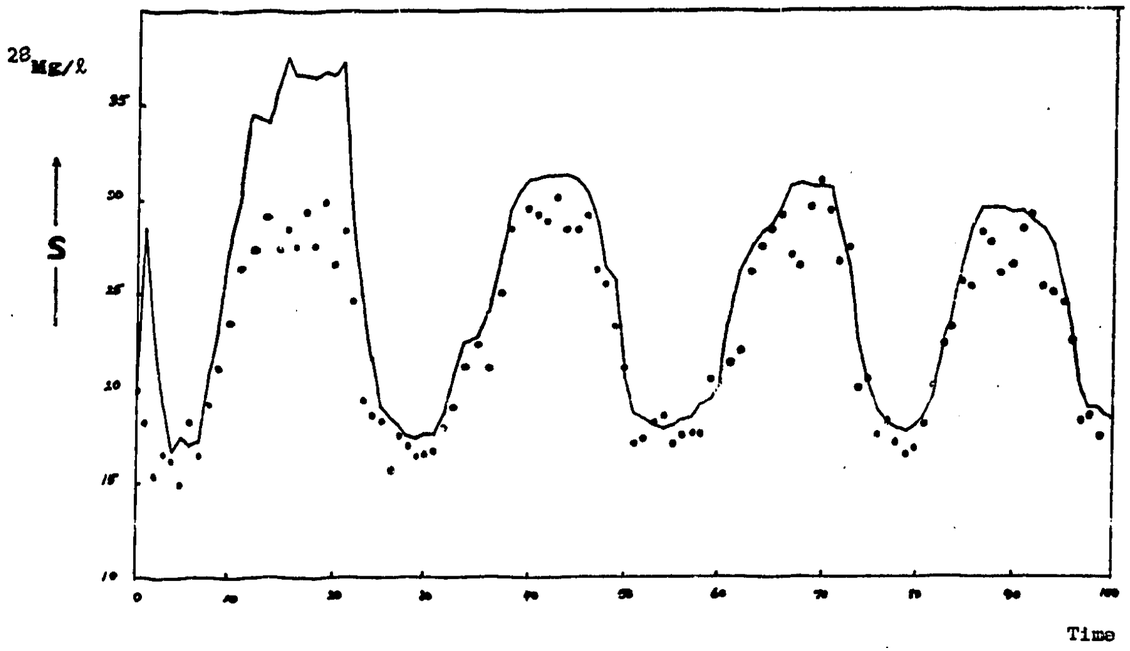
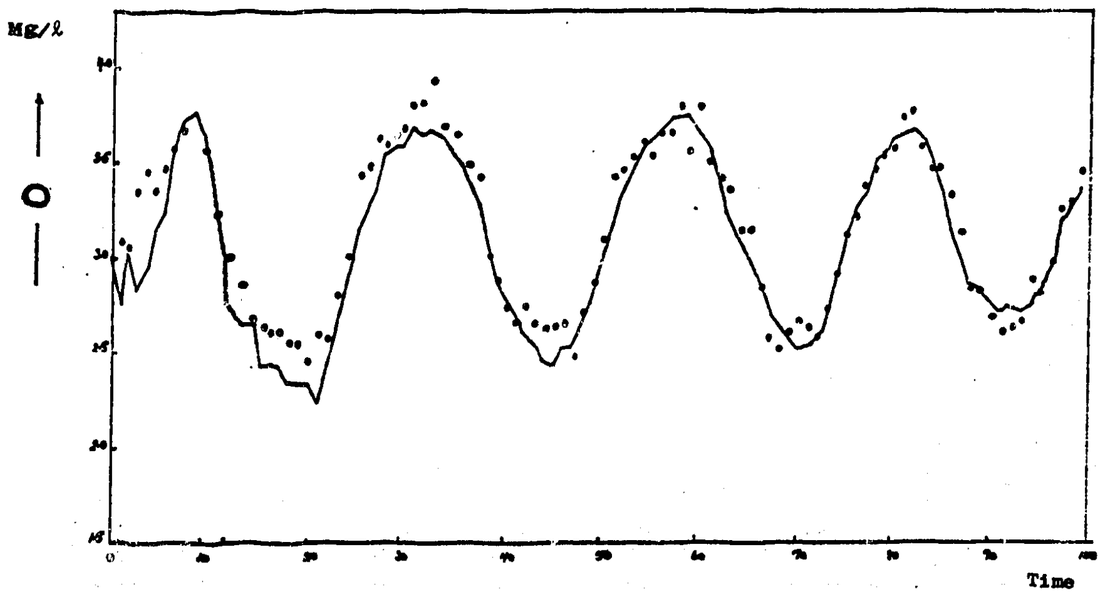


Figure 3.5a Observed (dots) and Reconstructed (line) Pollutant Concentration



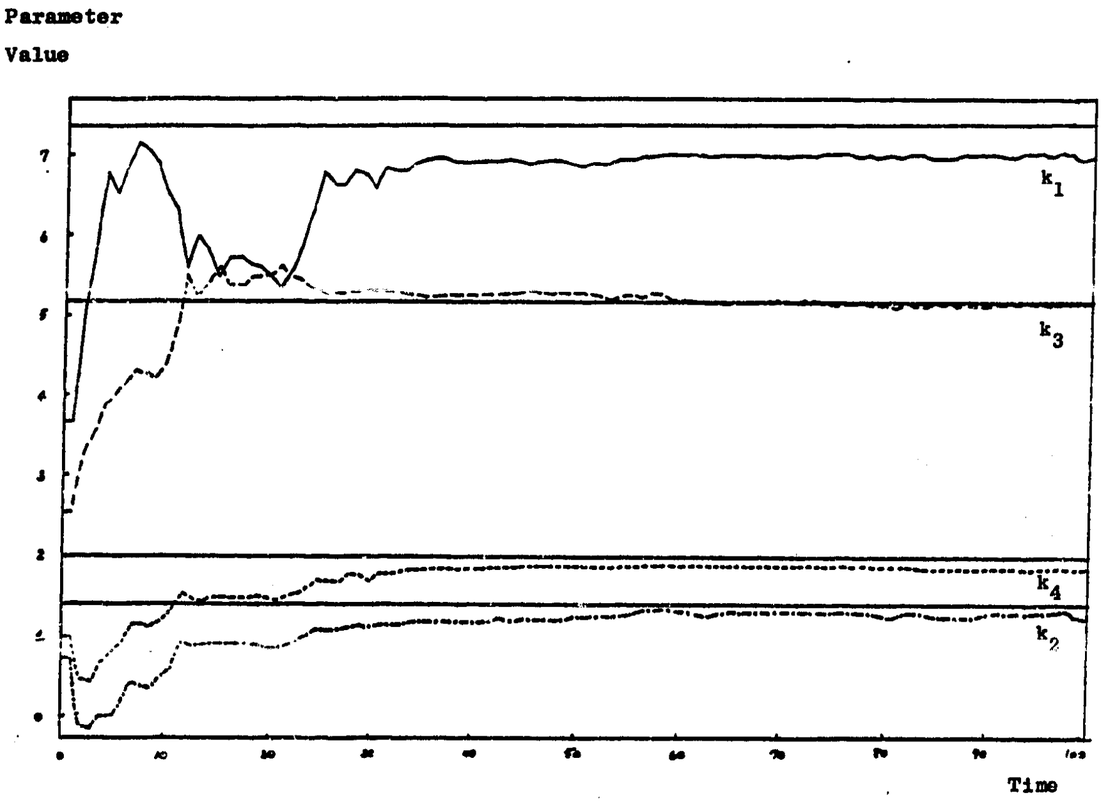


Figure 3.6 Parameter Estimates Showing Convergence

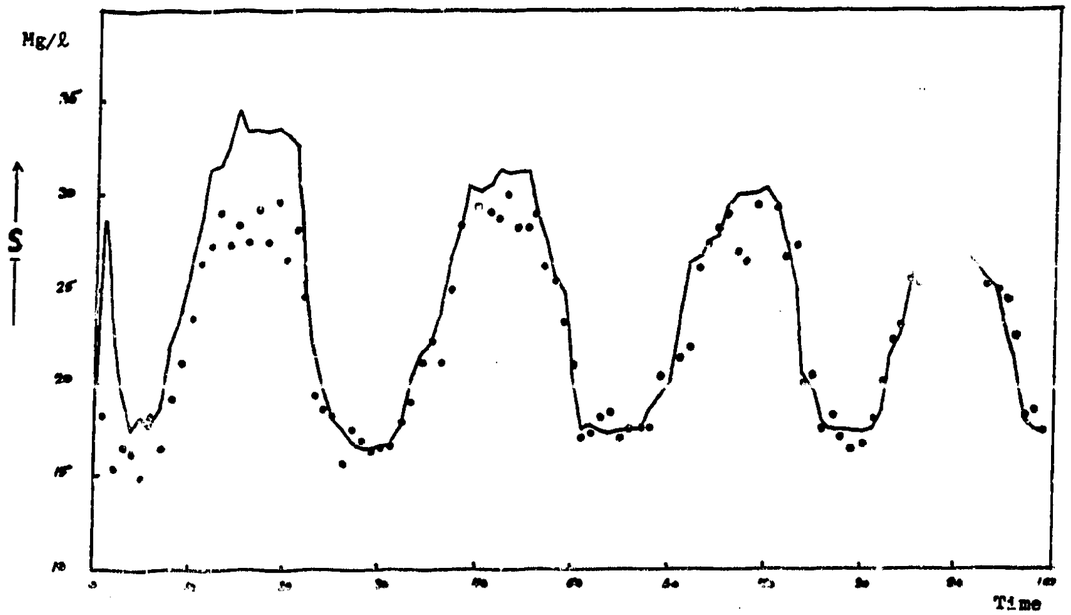


Figure 3.7a Observed (dots) Land Reconstructed Pollutant Concentration (line) assuming Random Walk in Parameters

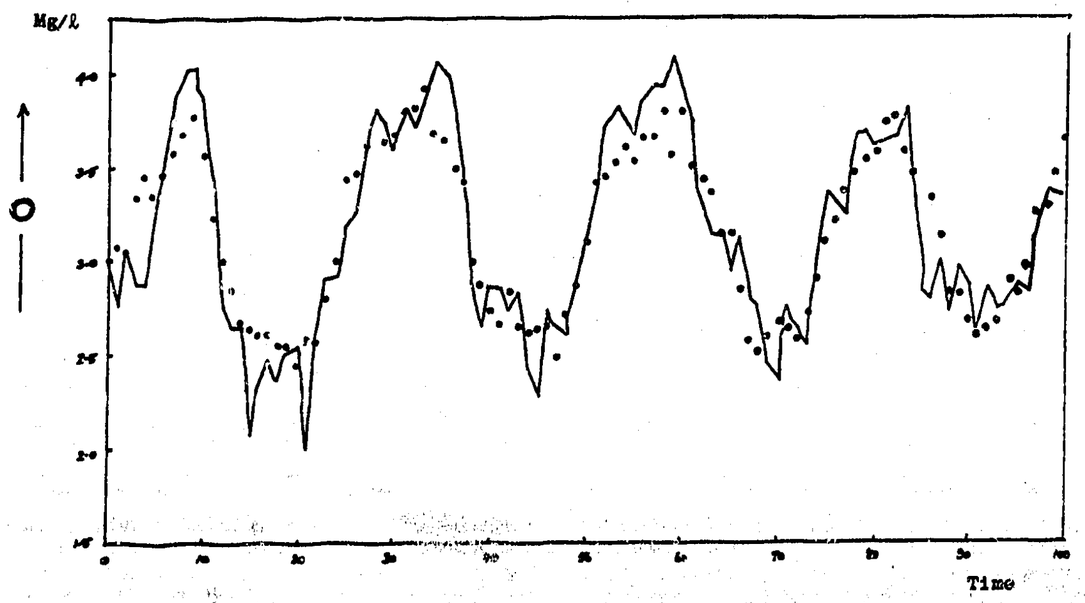


Figure 3.7b Observed (dots) and One Step Ahead Estimate of Oxygen Concentration (line) assuming Random Walk in Parameters

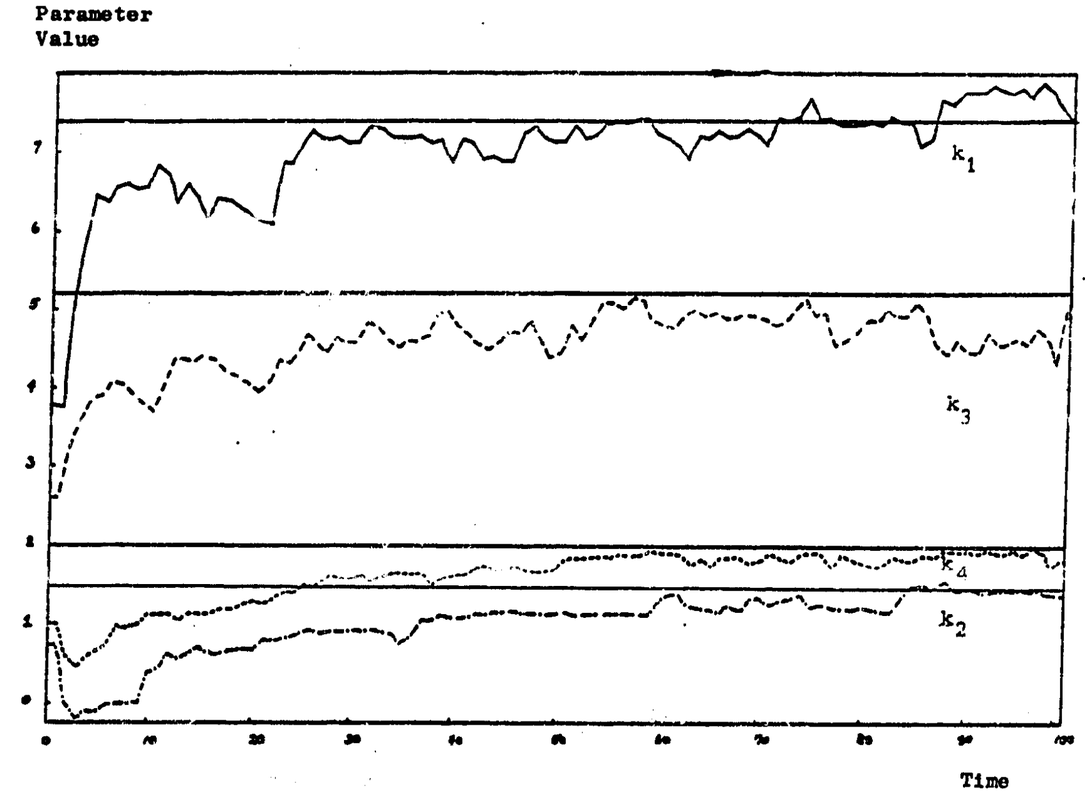


Figure 3.8 Parameter Estimates obtained with Random Walk Superimposed

system noise into the parameter dynamics a random walk structure is superimposed upon the parameters and this is seen to reduce the bias in the forecasts. However, an element of permanent scatter is introduced into the parameter estimates, as shown in Figure 3.8.

The objective of the study, that of combined state and parameter estimation in a non-linear model using the EKF, has been achieved with satisfactory results. The estimates of the unmeasured variable S (pollutant concentration) are well within observational tolerance and most of the error in the initial parameter estimates has been removed.

The use and capability of the EKF in real-time forecasting in non-linear water quality systems has been demonstrated and the incorporation of such estimation techniques into overall treatment control strategies remains a fruitful research area.

Computation aspects of the EKF are given in Appendix II together with brief user notes and a program listing. The EKF technique has been applied to river modelling problems during the workshop and these applications are described in Section 4 of the proceedings.

4. MATHEMATICAL MODELLING OF WATER QUALITY

4.1 Introduction

A common feature of many water quality studies is a nutrient budget to establish baseline knowledge about the system under consideration. Nutrients determine the trophic status of a river, reservoir or estuary, and nutrient information is required to gain an understanding of the ecosystem so that environmental problems can be foreseen and decisions made about its management and control. The important nutrients, Nitrogen (N) and Phosphorus (P), may be made available for plant growth by two mechanisms: the first is by external inputs such as effluent or field runoff and the second by recycling of nutrients already contained within the system. Considerable effort has been expended in attempts to measure the importance of various parts of the nitrogen and phosphorus cycles. Mathematical models play an important role in understanding the mechanisms of nutrient cycling and provide an integrated view of the nutrient system.

In Section 4.2, nutrient balance studies are described for two particular systems; a shallow estuary, the Peel-Harvey system in Western Australia, and the River Thames in the UK. In both studies detailed mathematical models have been developed which describe the dynamic behaviour of the systems. In the Peel-Harvey system the principal problem is the excessive growth of macroalgae, while in the case of the Thames Study, the transient violation of nitrate standards is of particular interest.

In addition to modelling nitrate in the Thames, a modelling study has been undertaken to investigate algal growth, death and transport along the river. Algae represent an important component in the nitrogen cycle and, in addition, create river management problems by blocking water filters and affecting water taste and smell. In Section 4.3 the generalised sensitivity analysis and EKF techniques described in Section 3 are employed in developing models of algal behaviour. The application of such techniques to biological systems is continued in Section 4.4 where a model of the activated sludge treatment process is developed.

4.2 Modelling nutrients in environmental systems

4.2.1 The Peel Inlet-Harvey Estuary system case study

4.2.1.1 History of the study

The Peel Inlet-Harvey Estuary system is a large (133 km²), shallow (average 1 m depth) waterway approximately 70 km south of Perth, Western Australia. Fig. 4.1 shows the location of the estuary and its principal hydrological features.

The excessive growth of green algae in the water of Peel Inlet and their accumulation and decay on the shores have created a nuisance for the last ten years. This has necessitated costly "cosmetic" measures to collect and remove weed accumulations near residential areas.

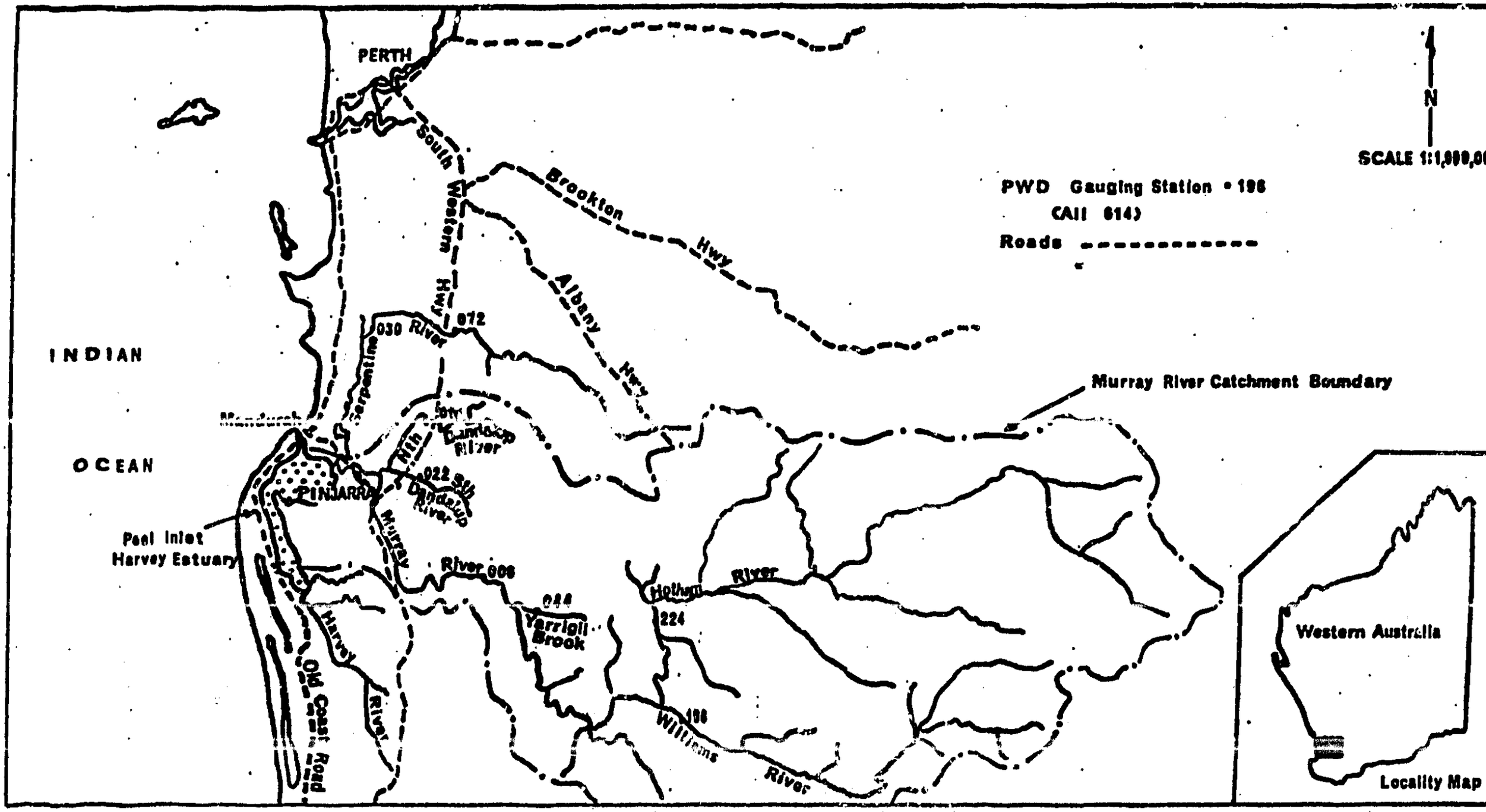


Figure 4.1 Location Map showing Peel Inlet - Harvey Estuary

Since 1976 a team of some twenty scientists from tertiary institutions and several government departments has studied all aspects of the Peel-Harvey estuarine system. The aim of the study was to determine the cause of the algal problem and to provide a basis for decisions about management of the estuary. A report was submitted in 1980 (Hodgkin *et al*, 1980).

4.2.1.2 Principal study areas

(a) Growth characteristics of the principal algal species, *Cladophora* aff. *albida*, and its nutrient requirements in relation to such environmental parameters as salinity, temperature, light and especially, the available nutrients. Mathematical modelling of *Cladophora* growth.

(b) Decomposition of the plants, storage of nutrients by surface sediments and their release in chemical form available to algae. Determination of the importance of this store of nutrients.

(c) Identification of external sources of nitrogen and phosphorus: quantifying these, their dispersion within the estuary, and loss to the sea. Mathematical modelling of historical data with respect to river flow and nutrient input.

(d) Hydrodynamics of the system and mathematical modelling of these.

(e) Sedimentology of the estuary in order to interpret its holocene history and to determine the organic content of the deeper sediments.

(f) Distribution and abundance of fish populations, with special reference to juvenile and non-commercial species, also crabs and prawns (commenced 1978). The estuary is not only a rich resource for both professional and amateur fishermen but, like other estuaries, it is also a nursery ground for marine as well as estuarine fish.

4.2.1.3 Initial study observations

(The following observations are attributable to Hodgkin pers. comm., 1980)

The cause of the algal problem is now well established: the excess of plant nutrients combined with the unusual growth characteristics of *Cladophora*. The plant grows as small cottonwool-like balls which normally lie on the bottom, but rise to the surface and drift to the shores. Decomposition occurs both onshore and in deeper water where the resultant black ooze and underlying sediments form a store of the principal plant nutrients, nitrogen and phosphorus. The input of nutrients to the estuary, phosphorus especially, has increased greatly over the last 25 years and it is eutrophic. The above statements are amplified in the following summary of relevant findings.

The shallow estuarine basin was flooded by the sea between 6,000 and 8,000 years ago, since when only 203 metres of sediment have accumulated, a slow rate of sedimentation. The rich and diverse fossil mollusc fauna of the lower sediments (dating from about 6,000 to 4,000 BC) indicates that the system was at first considerably more marine than it is now. The upper, more recent, sediments contain very few species of mollusc, species that still live in the estuary and tolerate the present extreme range of salinity (5-50 gm/l). Clearly the present restricted ventilation of the system is of relatively recent origin.

Although there are now only a few species of mollusc, worms, shrimps and other small invertebrates (the food of most fish), some of these are enormously abundant. Similarly, there are only a few species of aquatic plants and the present unbalanced condition of the ecosystem, associated with the great increase in nutrient input, favours excessive growth of new opportunistic species, especially *Cladophora* and planktonic algae.

In winter, the high nutrient concentrations that result from the brief period (about two months) of river flow favour planktonic algae, which are then abundant. However, low light and temperature levels inhibit growth of *Cladophora* and other bottom living plants at that time. In summer, when temperature and light are favourable to plant growth, nutrient concentrations are relatively low in the water, but the supply is supplemented by release of nutrients from the decomposing algae, the black ooze. Experiments and field observations have demonstrated that it is phosphorus rather than nitrogen that limits growth most of the time.

Analysis of data collected from 1949-56 in comparison with current data (1972-77) shows that input of phosphorus to the estuary has increased enormously over the past 25 years. There has probably also been a small increase in nitrogen input. About 90 per cent of phosphorus comes from coastal plain drainage; 60 per cent via the Harvey River, 20 per cent via the Serpentine River, and 10 per cent via the Murray River. The remaining 10 per cent is all that the hills catchment of the Murray River contributes.

Coastal plain soils are naturally deficient in phosphorus and there can be no doubt that most of the phosphorus in river water is derived from superphosphate. The application of this to coastal plain catchments has more than doubled in the last 30 years. Of some 1,380 tonnes of phosphorus applied in 1977-78 (as superphosphate), 103 tonnes (7.5 per cent) entered the estuary in drainage water.

It is evident that considerable quantities of nutrients are lost to the sea both during river flow and possibly by tidal exchange during no-flow periods. This loss is difficult to quantify so that retention rates can only be estimated on the basis of rather unsatisfactory data at this stage. Moreover this is made more difficult by an inadequate understanding of the mechanism of release of phosphate (the chemical form in which phosphorus is available to algae) from the sediment store of mineralised phosphorus. Experiments are continuing on this aspect.

4.2.1.4 The nutrient budget

As previously stated, external sources of nitrogen and phosphorus contribute large amounts of nutrients to the system via the three main river systems. Estimated nitrogen and phosphorus loadings for the water year 1977/78 are as tabulated below:

Source	N (tonnes)	P (tonnes)
Murray River	1,153	25
Serpentine River	116	22
Harvey River and drains	317	73
Rainfall	4	0.8
Total	1,590	121

The dispersion of these nutrients within the estuary and their loss to the sea via the narrow tidal channel are ill-defined. Nutrient concentrations have been determined on a weekly basis at a limited number of sites within the estuary and data on nutrient flux to the ocean are confined to two one-week intensive field exercises. The first of these provided little useful information as under the summer conditions of zero river flow, nutrient concentrations on both flood and ebb tides are very low.

However measurements taken at three hourly intervals for five days during late winter (August 1978) reveal a strong negative correlation between salinity and NO_2-N at the outer end of the entrance channel (A similar relationship is found to exist between salinity and NO_3-N within the estuary itself). These measurements were taken at Mandurah Bridge (see Figure 4.2) roughly in the main navigation channel of the estuary entrance. Every hour, on the hour, temperature, salinity and current speed and direction were taken at depth increments of 1 m (surface was taken to be at a depth of 0.5 m and bottom was approximately 5 m-5.5 m). Every 3 hours over a period of 96 hours water samples were taken, filtered and frozen for subsequent laboratory analysis for the following constituents:

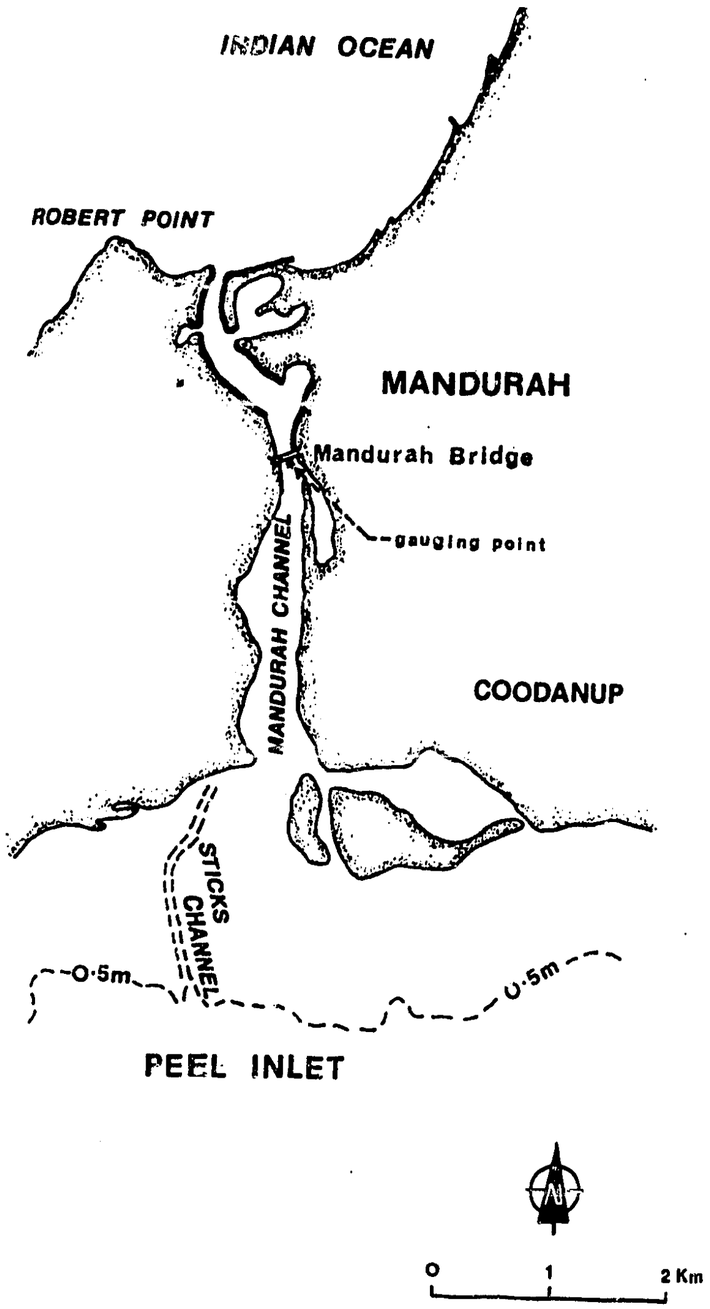


Figure 4.2 Mandurah Channel

Orthophosphate

Nitrate and Nitrite

Ammonia

Total N

Total P

Chlorophyll a

Plots of the limited time series (33 data points at 3-hourly intervals) of NO_3-N and salinity reveals a strong negative correlation between the two (see Figure 4.3) especially those measurements at depth (5 m) in the salt wedge. Between the measurement/sampling point on the Mandurah Bridge and the ocean, the entrance channel is 1.9 km long and the distance upstream of the Bridge to the estuary is approximately 3.6 km. At a representative tidal velocity of say 50 cm sec^{-1} the times to flush the entrance channel are 1 hour from the ocean to Mandurah Bridge and 3 hours from the estuary to the Bridge.

Both the salinity and NO_3-N data sets were analysed using the CAPTAIN package, a suite of time series analysis programs available at the Institute of Hydrology (Venn and Day, 1977) with the objective of establishing a relationship between the two series. Figure 4.4 shows the observed NO_3-N series and modelled NO_3-N output using input salinity series. The model is very simple (time delay = 0, MA = 1, AR = 0) but yields quite satisfactory agreement between observed and modelled NO_3-N.

For the Surface (0.5 m); $N_k - \bar{N} = -18.8\ (S_k - \bar{S})$

For the Bottom (5 m); $N_k - \bar{N} = -16.3\ (S_k - \bar{S})$

where N is NO_3-N in µg/L, $\bar{N}$ and $\bar{S}$ are mean nitrate and salinity levels.

and S is Salinity in ppt

These relationships can be used to determine nitrate levels from salinity measurements which can be easily measured over a long period of time. Nitrate [illegible] at the Mandurah channel site can be determined therefore and combined with other nitrate data for the estuary to compute an overall nitrogen budget. One problem in developing this budget is that water movement in and out of the estuary is dominated by low frequency barometrically induced oscillations with a periodicity of from 5 to 15 days. Indeed, this is so dominant that on occasions, the mean velocity does not even reverse sign with the diurnal tide. This means that in order to determine a nutrient budget, it will be necessary

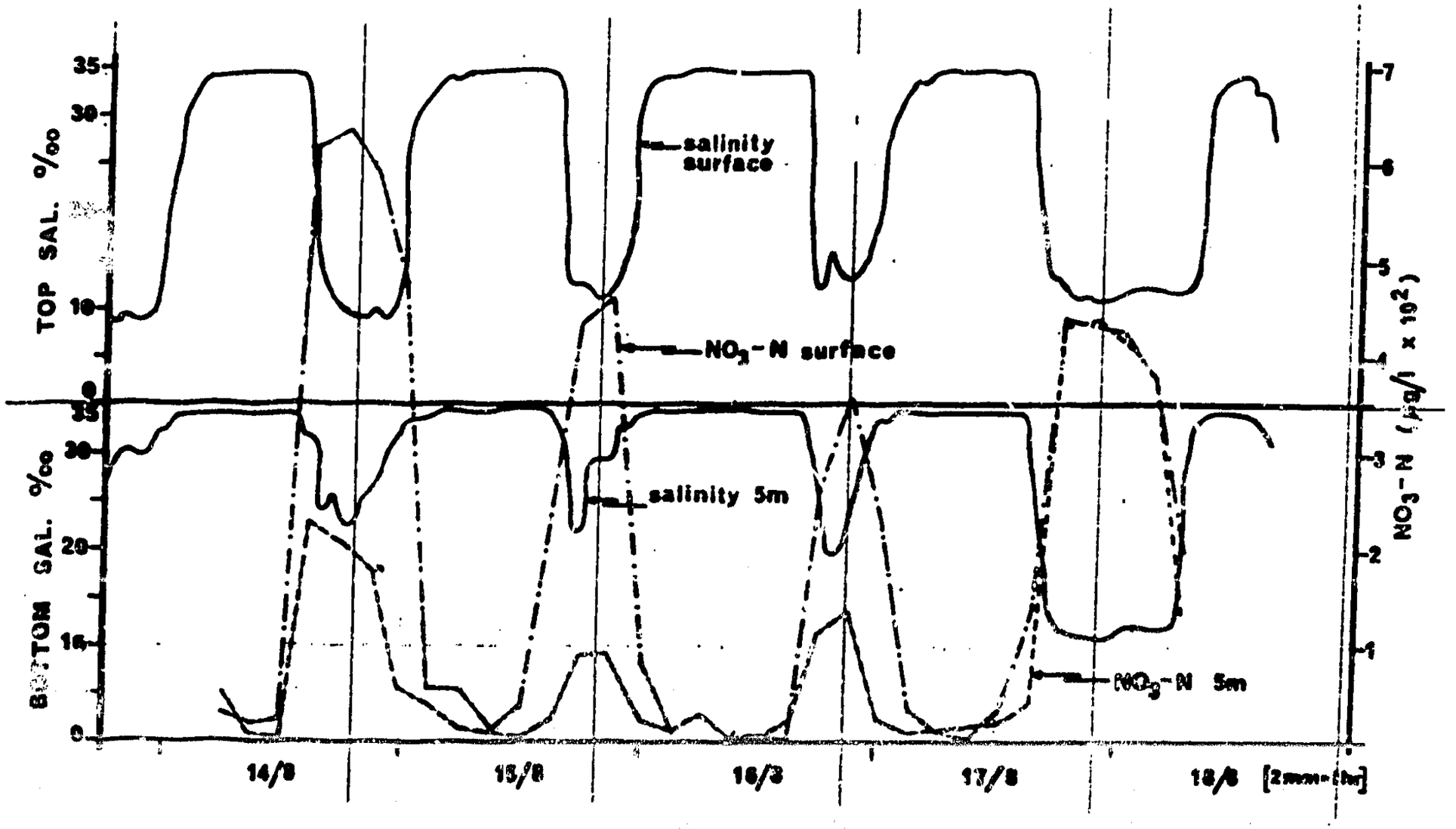


Figure 4.3 Salinity and Nitrate in Mandorah Channel

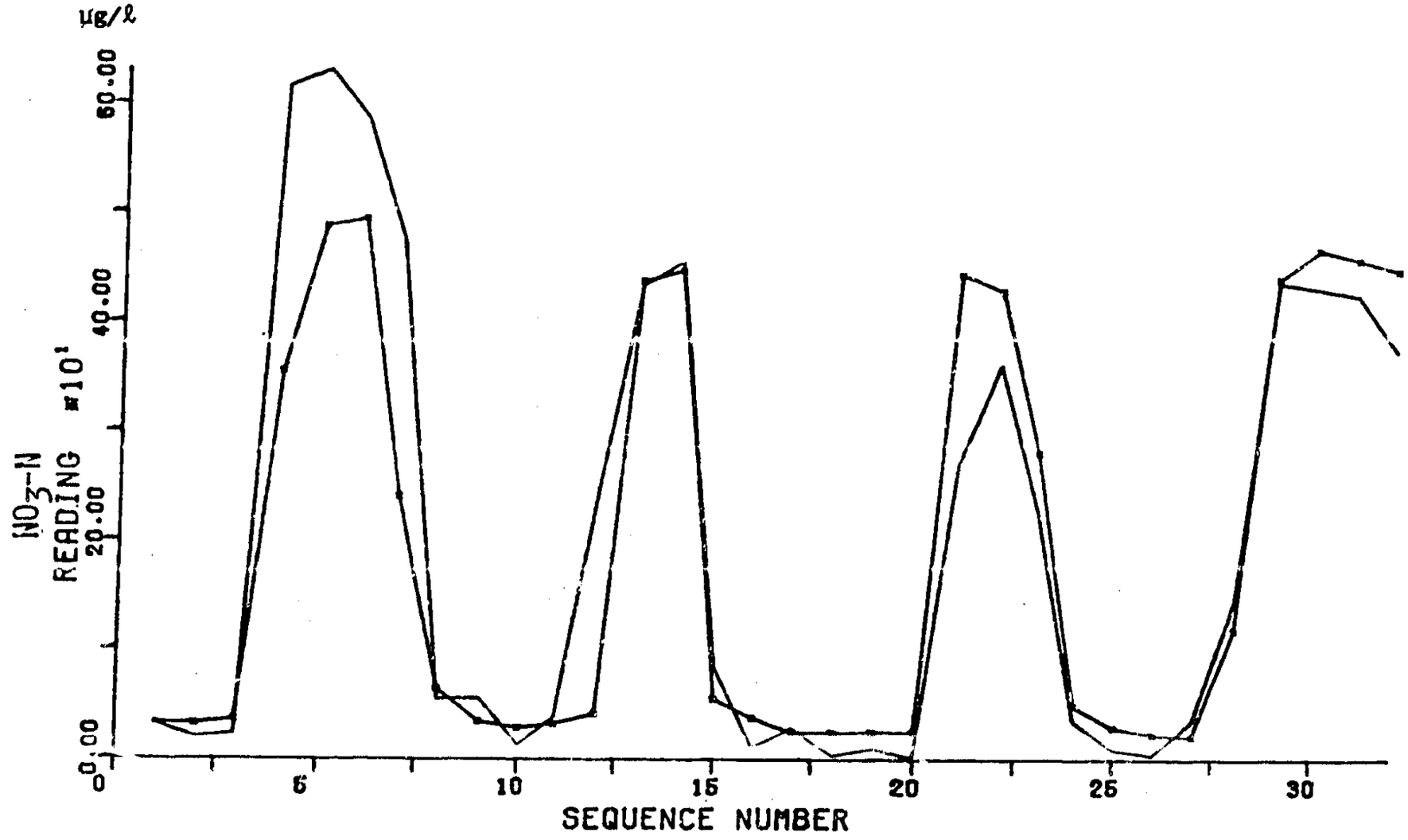


Figure 4.4a Observed (dotted line) and Simulated Nitrate at Surface

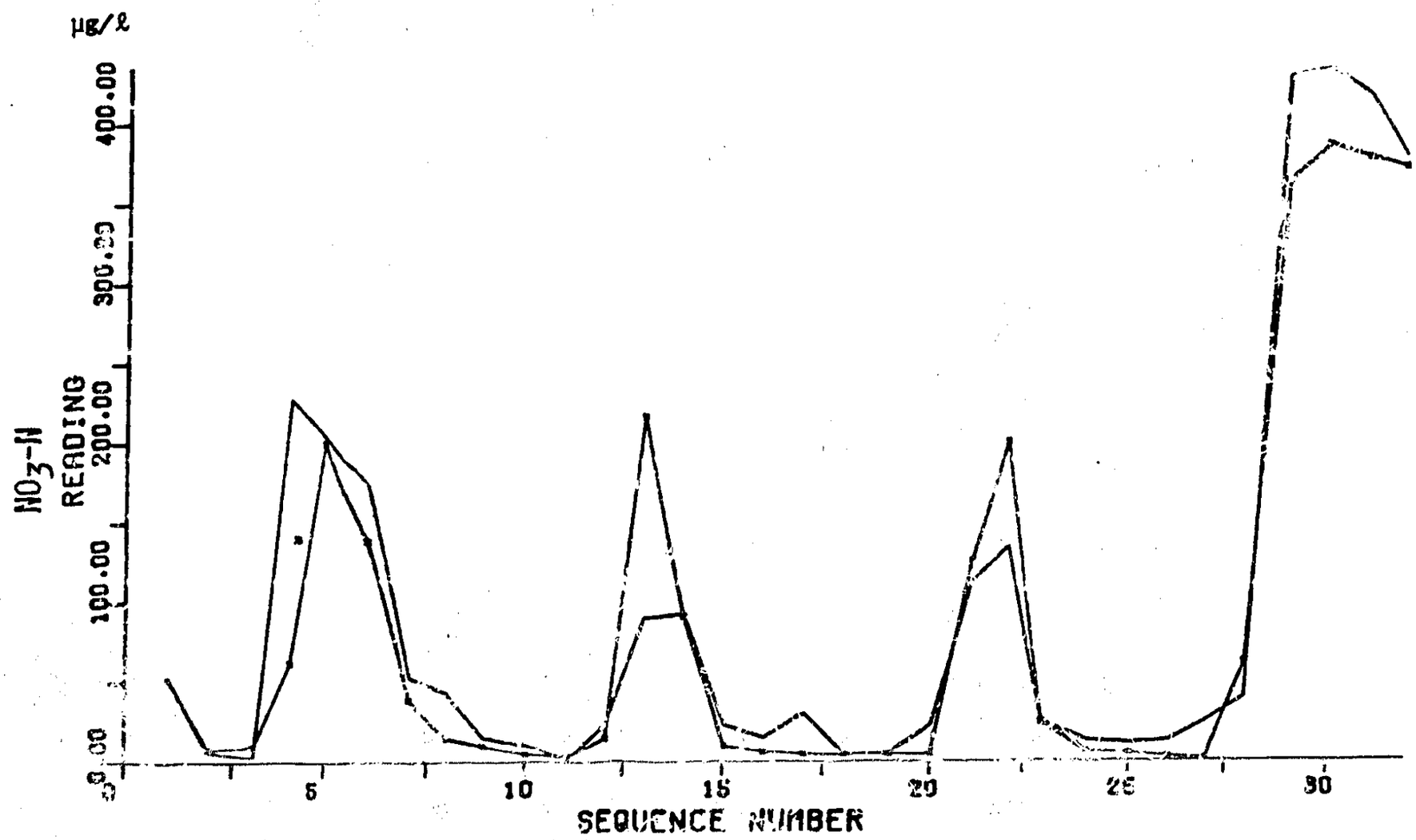


Figure 4.4b Observed (dotted line) and simulated Nitrate at 5 m.

to examine changes in nutrient load and concentration within the estuary itself, to evaluate losses and gains over a much longer period. This is possible using the weekly time series of nutrient concentrations at site within the system and a model of flushing developed previously (Humphries *et al*, 1980).

4.2.1.5 Computing an annual nutrient budget

Figure 4.5 is a flow chart of the weekly budget calculations made for each of two water years, 1977/1978 and 1978/1979. The observed weekly time series used were as follows (Hodgkin et al, 1980):

(a) River flow in the estuary ($m^3 x\ 10^3$) and river point nutrient concentrations ($\mu g.L^{-1}$) producing observed river input loads.

(b) Rainfall volume ($m^3 \times 10^3$) and average rainfall nutrient concentration ($\mu g.L^{-1}$) producing estimated rainfall nutrient loads directly onto the estuary water body.

(c) Estuary sector weighted water volume ($m^3 \times 10^3$) from tide heights continuously recorded and estuary water column nutrient concentrations measured at 7 sites within the estuary ($\mu g.L^{-1}$) producing estuary water column nutrient loads.

(d) Estuary sector flushing rate per week as calculated using the flushing model previously referred to (Humphries et al 1980).

As Figure 4.5 shows these are then used to generate an estimate of weekly nutrient loss or gain to the system, from which a forecast of the next week's water column load is made. The comparison of this load with that actually observed in the water column is a measure of the "errors" in the budget computation; ie uptake or release of nutrients from the estuary sediments, nitrogen fixation, marine nutrient input, as well as true measurement errors. Sediment uptake, for example would result in a negative value of observed minus forecast water column load.

Figure 4.6, the plot of the errors from this computation for total Nitrogen, shows a large negative "spike" at around week 48. This coincides with a massive N input of the major rivers in flood at that time and is an indication of an apparent large sediment uptake of nutrients.

An example of system generation of N would be a large positive value of observed minus forecast load. This can be seen in Figure 4.6 at about week 65 which coincides with an observed N-fixing algal bloom which occurred in the Harvey Estuary during November 1978. Similar computations were carried out for Phosphorus budgeting, but with rather less success (Hodgkin *et al*, 1980).

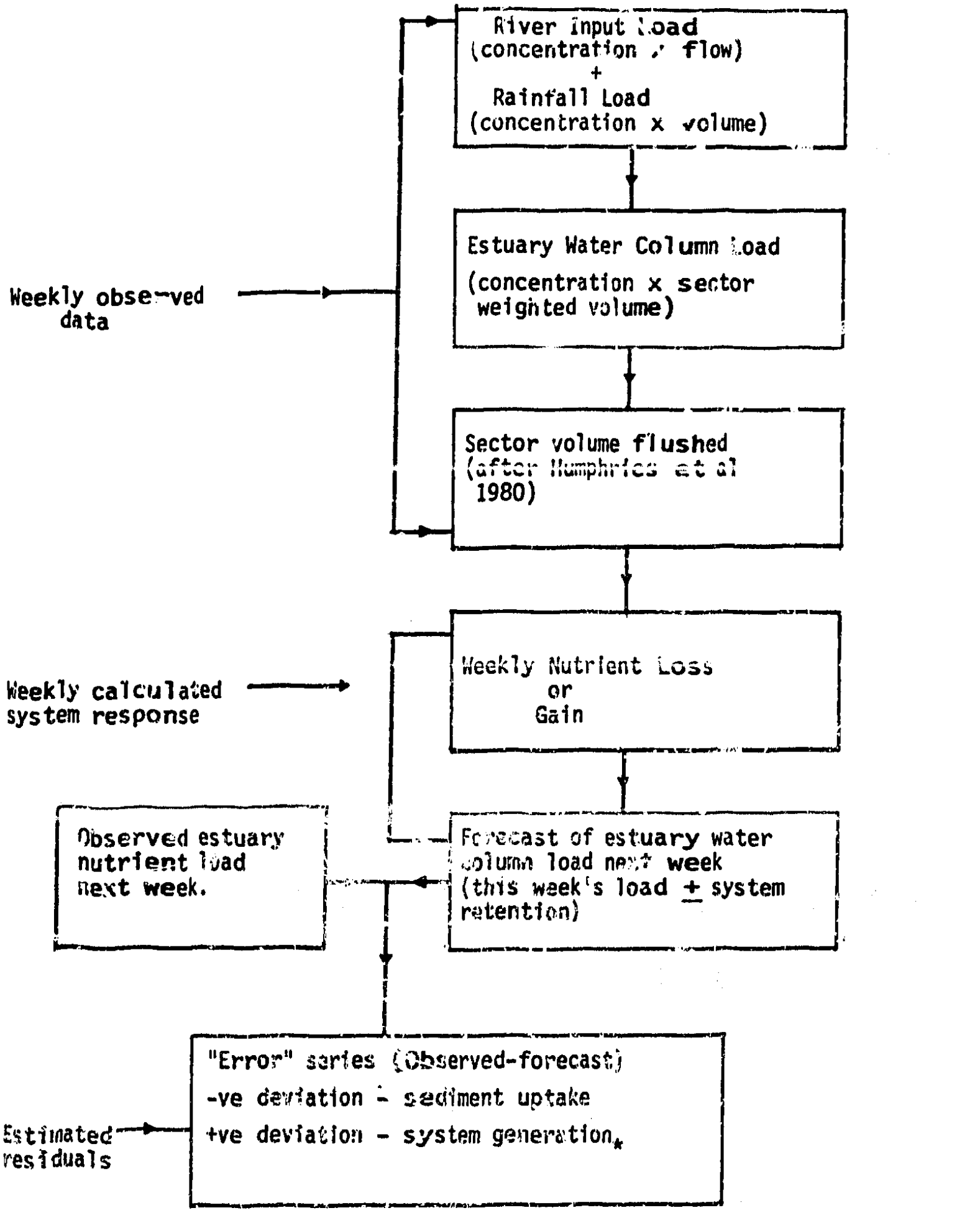
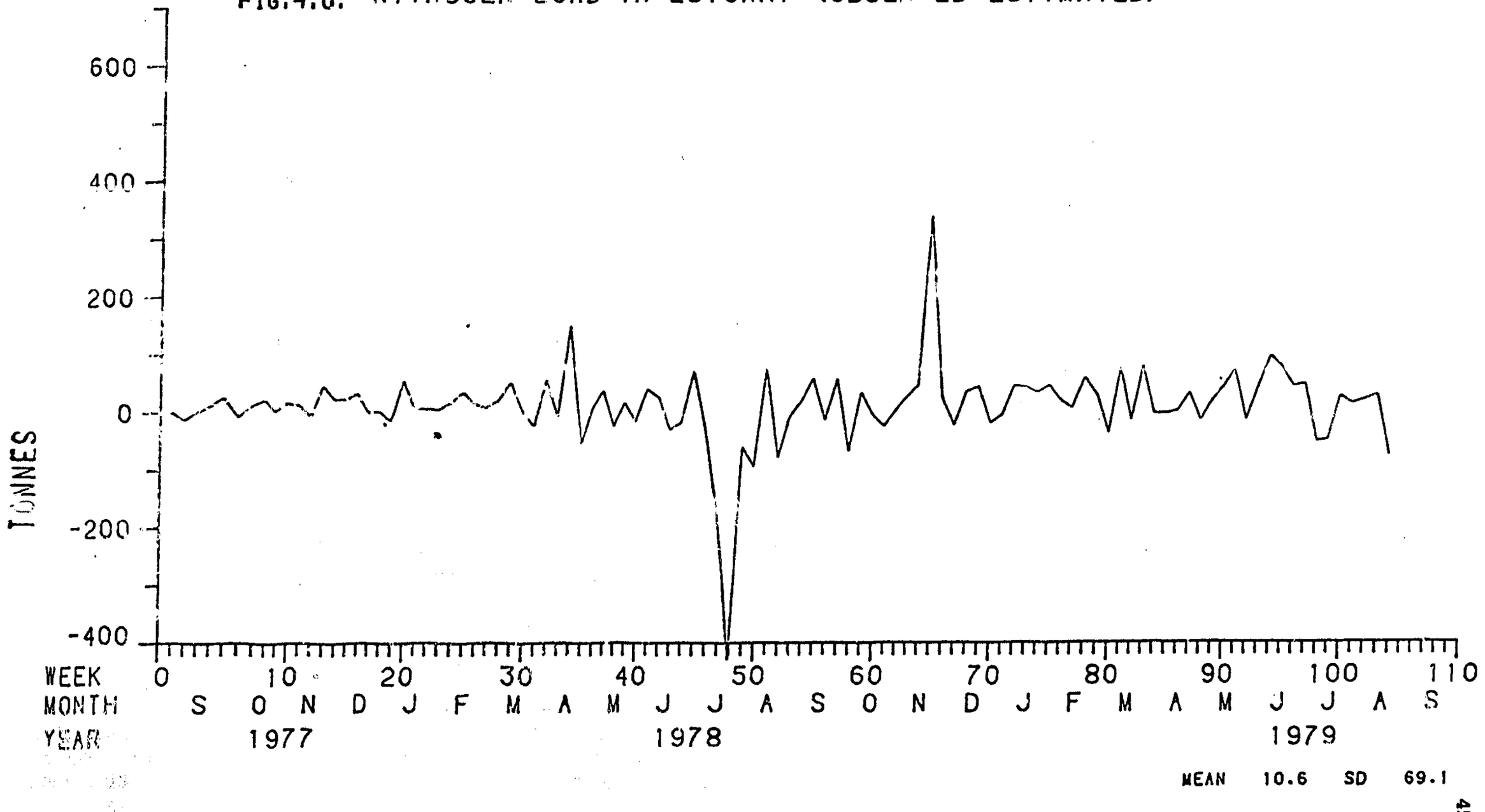


Fig. 4.5 : The Weekly Nutrient Budget Calculation (N and P).

FIG.4.6: NITROGEN LOAD IN ESTUARY (OBSERVED-ESTIMATED)



4.2.1.6 Control of the algae

The algal nuisance can be ameliorated by upgrading the present harvesting programme, but this would probably have to be continued indefinitely.

Long term control of the problem can best be achieved by reducing the present excess of nutrients available to the algae. Identification of practical measures by which this can be done obviously will require a lot of essential information on which control measures can be formulated. Decisions will have to be made on the basis of practicality - social, agricultural, engineering, political, and economic. At this stage the following appear to be the principal alternatives from the scientific viewpoint:

(a) Reduce the input of nutrients to the estuary by modifying fertilizer application techniques in such a way as to decrease release of phosphorus to coastal plain drainage;

(b) Reduce the internal supply of nutrients available to plants during the main growing season by removing the top layer of sediment (about 10 cm), together with living algae;

(c) Increase the loss of nutrients by, eg diverting major rivers direct to the sea, or by enlarging the Mandurah channel and thus speeding up river flow to the sea.

Summation of the weekly time series of N and P loads in the estuary enables an estimate of system gain or loss to be made. It was found that for the first year of the study (1977/1978), the estuary gained 1080 tonnes of N and 50 tonnes of P, graphically illustrating the build up of nutrients to the system responsible for the excessive algal growth. In the second, much drier year, where river flow was well below average, the estuary lost 344 tonnes N but still gained 24 tonnes P. This seems to indicate that the problem will continue, even if river flow continues to be well below the long term average, as has been the case for some years.

4.2.2 The Thames nitrate study

4.2.2.1 Introduction

Although, in general, water quality in the UK has been improving in recent years, the situation with regard to certain variables such as nitrate has deteriorated. Green (1978) reports that a significant upward trend in nitrate concentrations is evident in many groundwater and surface-water abstractions used for public supply purposes in the Anglian region; current information appears to indicate that these increases may be primarily associated with the increasing intensity of, and/or improvements in arable farming in the region in the last twenty years. Approximately 50 per cent by volume of the Anglian Water Authority's abstractions have consistently exceeded the WHO/EEC standard of 11.3 mg/l nitrate-N with a small number of groundwater sources in strategic locations

approaching 22.6 mg/l; the majority of surface abstractions, both direct river and reservoir, have exhibited concentrations in excess of 11.3 mg/l nitrate-N for transient periods (days to months).

In the Thames River System similar problems have arisen. For example, the mean annual nitrate concentration rose from 4.2 mg/l Nitrate-N in 1968 to 7.7 mg/l Nitrate-N in 1979 at Walton, the intake for the lower Thames reservoirs serving London. In the winter of 1973/74 the concentration was above the EEC limit for two weeks and in 1976/77 the limit was exceeded for four weeks. In order to devise a regional strategy to manage the nitrate problem Thames Water Authority have initiated a mathematical modelling study.

4.2.2.2 Preliminary problem identification studies

The initial approach taken in the Thames nitrate study has been to analyse the historical river nitrate series by relating it to a derived series of nitrogen inputs from diffuse and point sources (Onstad and Blake, 1980). Annual inputs from agriculture for the period 1921-75 were estimated from published statistics on county land use, animal numbers, fertiliser use and crop yield, and inputs from sewage effluents were estimated from population statistics. The study indicated that changes in agricultural inputs could account for 78% of the variance of nitrate at Walton and this figure compared with 80% estimated for the Bedford Ouse River by Owens *et al* (1972). A mathematical relationship between the input nitrate loads and the river concentrations for the Thames has been developed by Onstad and Blake and from the analysis two principal paths identified for soil nitrate movement. These are a rapid path through the soil layer only, and a slower path through the underlying aquifers. The main features which helped to establish this relationship were the release of soil nitrogen in the years 1939-43 after the ploughing of grassland and the increased fertiliser applications from the early 1960's. Having identified and estimated a time series model relating nitrate load to river-nitrate level, the model was used to predict future mean nitrate concentrations given a range of possible futures for agriculture. The analysis suggested that the surface sources in the Thames system are vulnerable and, given that there is little spare capacity in London's water resources at present, a strategy is needed for the design and management of the Thames water resource system from a nitrates standpoint.

Several options are available for managing the nitrate problem such as the artificial denitrification of water, the provision of additional reservoir storage to promote natural denitrification, the blending of river water with lower-nitrate groundwater or reservoir water, the denitrification of sewage effluents, and the more efficient operation of groups of reservoirs to increase the store of low nitrate water. These options have to be considered in association with those for a new resource in the Thames Region. It has been decided that an additional resource is needed for London irrespective of quality constraints and options range from new reservoirs to enlarged groundwater schemes. Each of those options affect water quantity and quality and an integrated mathematical model would therefore be required to evaluate the complex range of options.

A water quantity mass balance model which has been developed by the Thames Water Authority for resource evaluations provides hydrological inputs for a quality model of the Thames system. The quantity model consists of a number of sub-models representing components of the water resource system, namely the soil zone, aquifer, rivers and reservoirs components. Nitrate models are needed for each of these components as well as sewage treatment works and water denitrification plants. In order to obtain a quality model within a reasonable time scale, component models have been developed in a collaborative study between the Water Research Centre, the Institute of Hydrology and the Thames Water Authority. Soil zone and aquifer components have been developed at the Stevenage and Medmenham Laboratories of the Water Research Centre (WRC) and the reservoir models by the Thames Water Authority. The Institute of Hydrology has been responsible for developing a generalised flow and quality model for the multi-tributary, multi-reach system which can accept inputs from groundwater, surface runoff, tributary inflows, sewage discharges and can also account for water abstractions. Figure 4.7 shows the reach structure necessary for the Thames and Lea model with reach boundaries defined by the locations of inflows and abstractions.

An important objective of the water quality model is to evaluate the risks of nitrate exceeding EEC standards at key locations given a specified strategy. Information in the form of probability distributions is therefore required so that, for example, the number of days nitrate is above the standard can be estimated. To provide information on the day to day changes in nitrate concentration, a dynamic flow and quality model is required of the type developed for the Bedford Ouse Study (Whitehead *et al*, 1979, 81).

4.2.2.3 Dynamic modelling of nitrogen in rivers

Nitrogen transformations in a river such as the Thames are complex as indicated by Figure 4.8. The principal processes involve the nitrification of ammonia to nitrite and hence to nitrate under aerobic conditions and the reduction of nitrate to nitrogen or ammonia under anaerobic conditions. Plants and algae also assimilate nitrate from the water column during growth periods and this nitrogen is recycled eventually during the decay of cellular organic matter.

Most modelling studies have been restricted to the nitrification process since the discharge of ammonia into rivers from industrial or domestic effluents has a significant effect on fisheries, water supply and, as indicated in the following equations, dissolved oxygen levels.

$$NH_4^+ + (3/2)O_2 \xrightarrow{\text{Nitrosomonas}} NO_2^- + 2H^+ H_2O \qquad (4.1)$$

$$NO_2^- + (\tfrac{1}{2})O_2 \xrightarrow{\text{Nitrobacter}} NO_3^- \qquad (4.2)$$

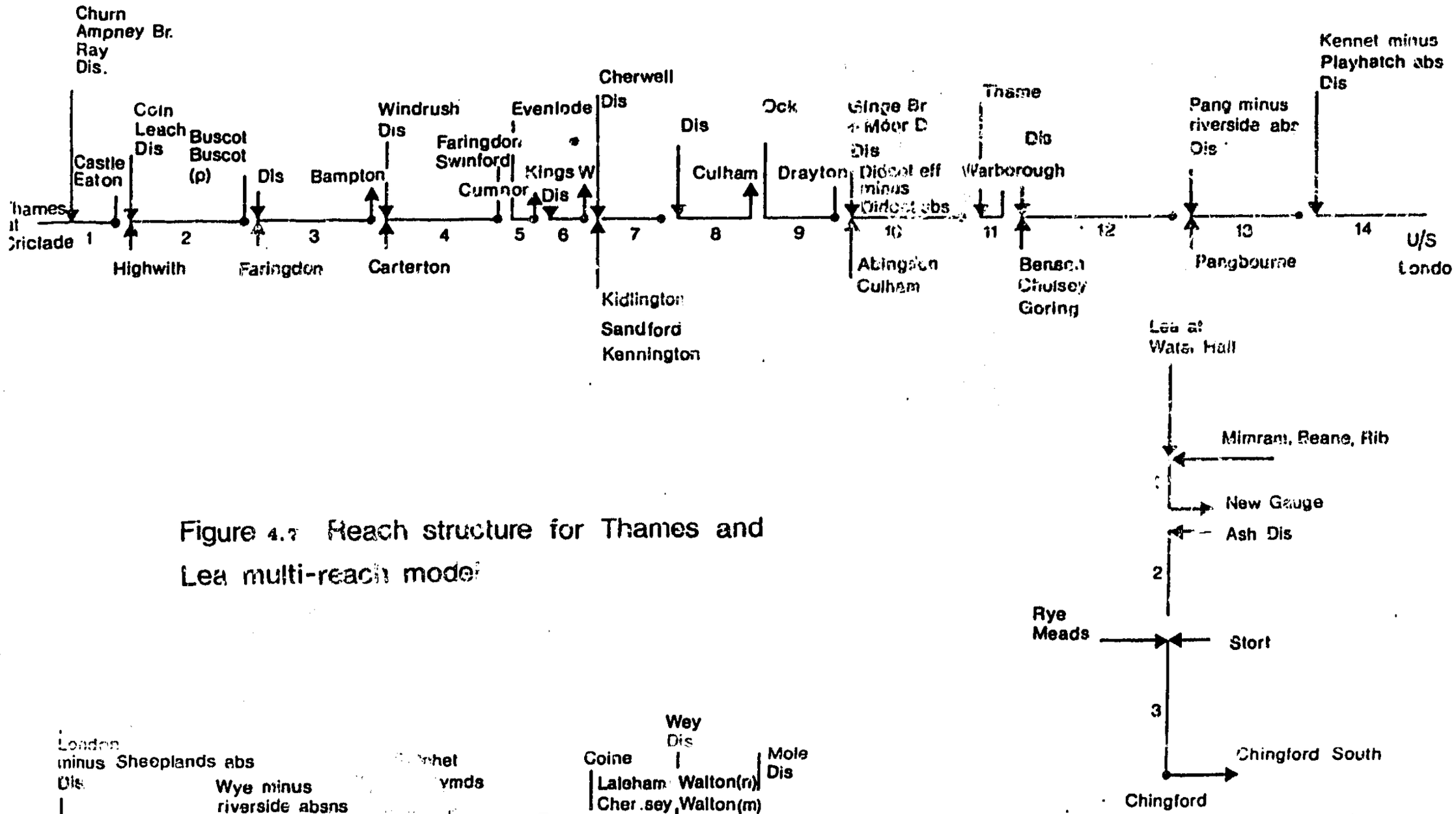
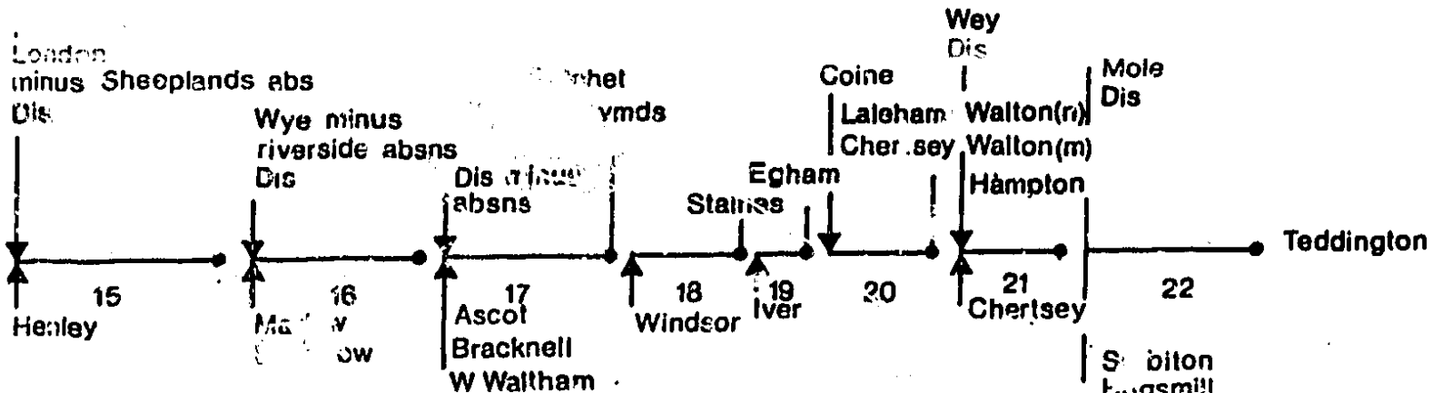


Figure 4.7 Reach structure for Thames and Lea multi-reach model



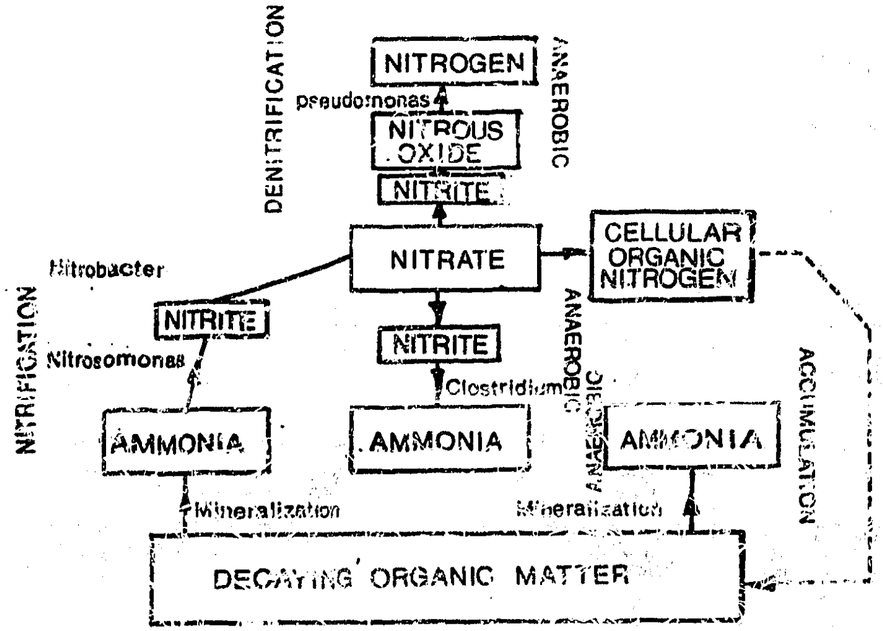


Figure 4.8 Nitrogen Transformations in River Systems

Nitrosomonas and nitrobacter are autotrophic bacteria responsible for the oxidation process. Since this process is a slow stage biochemical reaction most mathematical models that have been developed (O'Connor et al, 1976) are in the form of a set of coupled partial differential equations such as;

$$\frac{\partial N_1}{\partial t} = - u\frac{\partial N_1}{\partial x} + D\frac{\partial^2 N_1}{\partial x^2} - k_1 N_1 \qquad (4.3)$$

$$\frac{\partial N_2}{\partial t} = - u\frac{\partial N_2}{\partial x} + D\frac{\partial^2 N_2}{\partial x^2} + k_1 N_1 - k_2 N_2 \qquad (4.4)$$

$$\frac{\partial N_3}{\partial t} = - u\frac{\partial N_3}{\partial x} + D\frac{\partial^2 N_3}{\partial x^2} + k_2 N_2 \qquad (4.5)$$

where N_1, N_2 and N_3 represent ammonia (as N), nitrite (as N) and nitrate (as N) concentrations, u represents velocity, D is the dispersion coefficient, x represents distance along the river, t is time and K_1 and K_2 are reaction rates. These equations are often simplified to first order lumped parameter models where advection forces dominate.

Surprisingly there has been relatively little modelling research on the nitrate reduction process in rivers. The process of denitrification is represented by the following equation:

$$6\ NO_3^- + 5\ CH_3OH \rightarrow 3\ N_2 + 5CO_2 + 7\ H_2O + 6OH^- \qquad (4.6)$$

Biological denitrification is promoted by a large number of bacteria which contain nitrate reductases, enzymes that mediate the reaction. The nitrogen gas formed by this reaction may be transferred from the water to the atmosphere if the nitrogen concentration exceeds the saturation concentration. The reduction processes occur in the mud or at the mud/water interface and as indicated in equation (4.6) the reaction requires an organic carbon substrate. Because of the complexity of the reaction mechanisms it is normally assumed that the reaction kinetics are first order (O'Connor, 1976) with the transport equation for nitrate written as:

$$\frac{\partial N_3}{\partial t} = - u\frac{\partial N_3}{\partial t} + D\frac{\partial^2 N_3}{\partial x^2} + k_3 N_2 - k_4 N_3 \qquad (4.7)$$

where k_4 is the denitrification rate.

Empirical evidence for the first order reaction kinetics is given by Toms et al. (1975) but with the modification that the reaction rate is a function of bed surface area and temperature. In addition there is considerable spatial variation in the denitrifying capacity of sediments in lowland eutrophic rivers such as the Lea and the Thames and selecting a representative denitrification rate is problematical. In the Bedford Ouse Study (Whitehead *et al*, 1981) the denitrification rate was estimated using the extended Kalman filter for a short reach of river. However, daily nitrate data were available for the latter study and such a data base is not available for the Thames and Lea system. Nonetheless, sufficient data are available for preliminary analysis and modelling.

4.2.2.4 The Thames River Stream flow Model

In order to model nitrate it is necessary to simulate streamflow in all the reaches of the river shown in Figure 4.7. In the Thames streamflow model each reach is characterised by a number of cells and the model for flow variations in each cell is based on an analogy with the lumped parameter equations for the variations in concentration of a conservative pollutant under the assumption of uniform mixing over the cell (Whitehead et al, 1979). The model may be viewed in hydrological flow routing terms as one in which the relationship between inflow I, outflow, Q, and storage, S, in each cell is represented by the continuity equation:

$$\frac{dS}{dt} = I - Q \qquad (4.8)$$

with

$$S = \tau Q$$

where τ is a travel time parameter. If τ is fixed, then this does not allow any variation in travel time with flow; to achieve this, τ is expressed as

$$\tau(Q) = \frac{L}{UN} \qquad (4.9)$$

where N is the number of compartments in the reach, L is the reach length and U, the mean flow velocity in the reach, is related to discharge through

$$U = a Q^b \qquad (4.10)$$

where a and b are coefficients to be estimated. The resulting differential equation to be solved is then

$$\frac{dQ}{dt} = \frac{aN}{L(1-b)} Q^b (I - Q) \qquad (4.11)$$

The value of N affects the relative importance of floodwave advection and dispersion in a reach, values of N, a and b can be determined by calibration on an observed record of downstream flow or from tracer experiments (see Whitehead et al, 1974).

Given information on upstream and tributary inputs, the flow routing model can be used to derive simulations of downstream flow by solving the differential equation (4.11). The equation is solved using a numerical integration technique which contains an automatic adjustment to the integration step length. This is particularly useful since under periods of low flow and high residence times, the integration step length can be increased thereby saving computer time. Under high flow conditions, however, residence times are reduced and in order to solve the equation to the same accuracy, it is necessary to reduce the integration step length. Since this is achieved automatically, there are relatively few numerical integration problems. Figure 4.9 shows simulated flow compared to observations for the years 1971, 75 and 76; 94% of the variance is explained and the model provides a sound basis for subsequent water quality studies.

4.2.2.5 The Thames River Nitrate Model

The dynamic water quality model for the Thames is based on a mass balance approach for a non-conservative variable and can be written for a reach cell as:

$$\frac{dx(t)}{dt} = \frac{Q_i(t)}{V_e} u_i(t) - \frac{Q_o(t)}{V_e} x(t) + \frac{Q_G(t)}{V_e} u_G(t) + \frac{Q_T(t)}{V_e} u_T(t) + \frac{Q_S(t)}{V_e} U_s(t) + \frac{Q_E(t)}{V_e} u_E(t) - \frac{Q_A(t)}{V_e} u_A(t) - \frac{N}{d} (10^{0.0283\theta}) x(t) \qquad (4.12)$$

where x(t) is the output nitrate concentration, u(t) is the input nitrate concentration, $Q_i(t)$ and $Q_o(t)$ are upstream and downstream flow rates, K is the denitrification rate, d is a river depth, θ is temperature and the subscripts G, S, T, E and A refer to groundwater inflows, surface runoff, tributaries, effluents and abstractions, respectively. Inputs from groundwater, daily surface runoffs and tributaries are defined by the WRC soil zone and groundwater models. V_e represents the 'effective' volume of the cell; this allows for short circuiting effects in the river and presence of dead zones, and can be determined by the relationship

$$V_e = \tau Q_o \qquad (4.13)$$

where τ is the mean residence time in the reach cell defined previously in the flow model. The effective volume may also be determined directly from a dye tracer experiment.

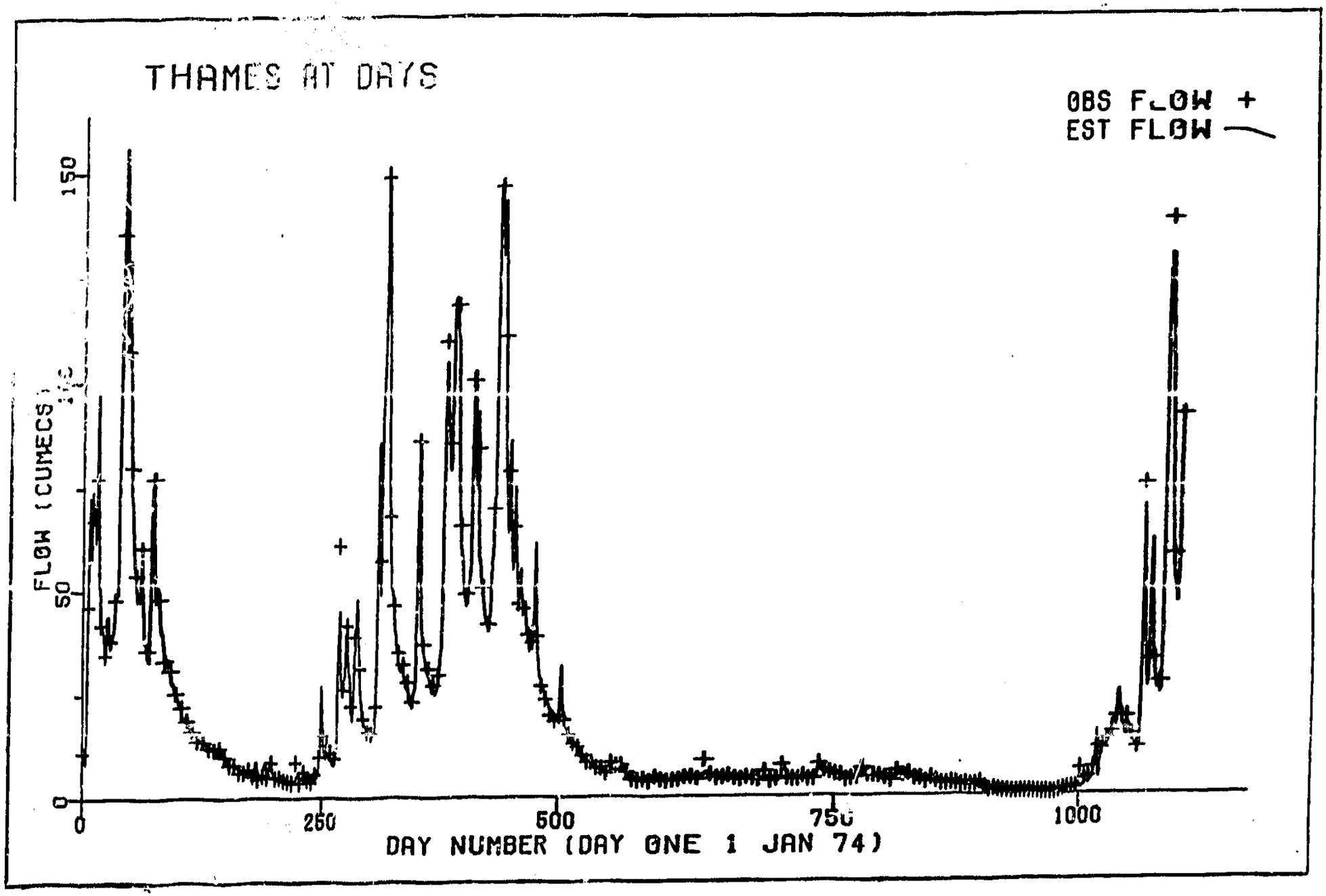


Figure 4.9 Observed and simulated daily flows at Day's Weir for 1974, 75 and 76.

The river model consists of a multi-reach structure with a total of 22 reaches for the River Thames and 3 reaches for the River Lea (see Figure 4.7). Each reach model consists of a number of cells each described by equation (4.12) and the total 'cells in series' model can be shown theoretically and experimentally to be equivalent to the advection-diffusion equation description (Whitehead et al, 1979, 1981). Numerical solution of the diffusion equation provides similar results to the cells in series model since in both cases the river is sub-divided into finite, well-mixed volumes and concentrations are computed as discrete values. Whereas numerical dispersion obtained when solving the partial differential equation is an undesirable by-product, in the cells in series model the dispersion due to discretisation is used to simulate real dispersion. In addition the cells in series model produces skewness without additional terms as in the transport equation and retention of material in the cells is an integral part of the model formulation.

The Thames and Lea nitrate model must await the complete development of the WRC soil and groundwater model for final validation. However, preliminary model testing has been possible using nitrate data obtained by routine sampling of the Thames and the Lea. In low flow summer conditions the long residence times associated with the reaches allow for a significant loss of nitrate by denitrification as indicated in Figures 4.10 and 4.11. In addition low temperatures in winter reduce the denitrification rate. At this stage the Thames simulation shows a significant discrepancy during the winter months. This is due to the short term variations in surface runoff under winter storm events which are often missed by routine sampling. The WRC soil model should provide information on these inputs and allow the detailed estimation of denitrification rates. Initially k/d in equation (4.12) has been set to 0.05 days^{-1}, an average figure obtained from studies by WRC, Thames Water Authority and the Bedford Ouse Study.

4.2.2.6 Preliminary model applications

Using the model it has been possible to investigate the percentage nitrate loss by denitrification at different locations along the river. The system was simulated initially with K in equation (4.12) set to zero, ie no denitrification. This gave a mass nitrate load at each reach boundary and this can be compared with the nitrate load allowing for denitrification processes. The total percentage loss at four locations on the river for average summer and winter conditions is given in Table 4.1. Several interesting features are illustrated in these results. Firstly, as might be expected from temperature differences, the summer nitrate losses are significantly higher than the winter losses. In 1976, for both summer and winter conditions, nitrate loss is higher than 1974 and 1975 and this is probably due to the very low flow rates through 1976 (see Figure 4.9). Of particular interest is the high denitrification occurring in the river between Swinford and Caversham. The reason for this high percentage loss is the considerable length of this stretch of river and the long residence times available for denitrification. Another possible factor is the growth of algae along this reach of the river. Caversham chlorophyll levels are a factor of four times higher than levels at Swinford and uptake of nitrogen by algae and plants can contribute significantly to the water

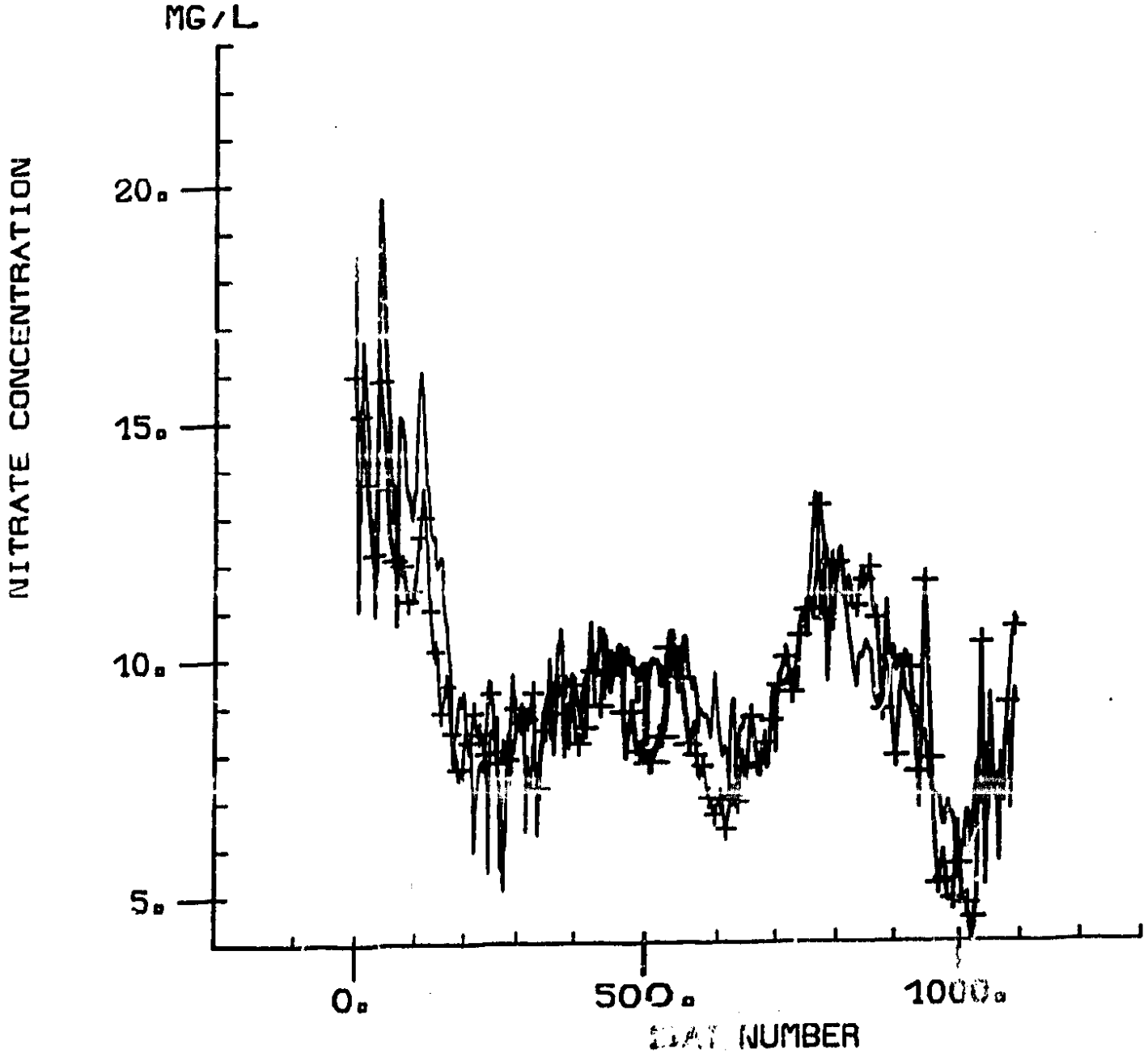


Figure 4.10 Simulated and Observed Nitrate (NO_3 as N) concentrations on River Lea 1974-76

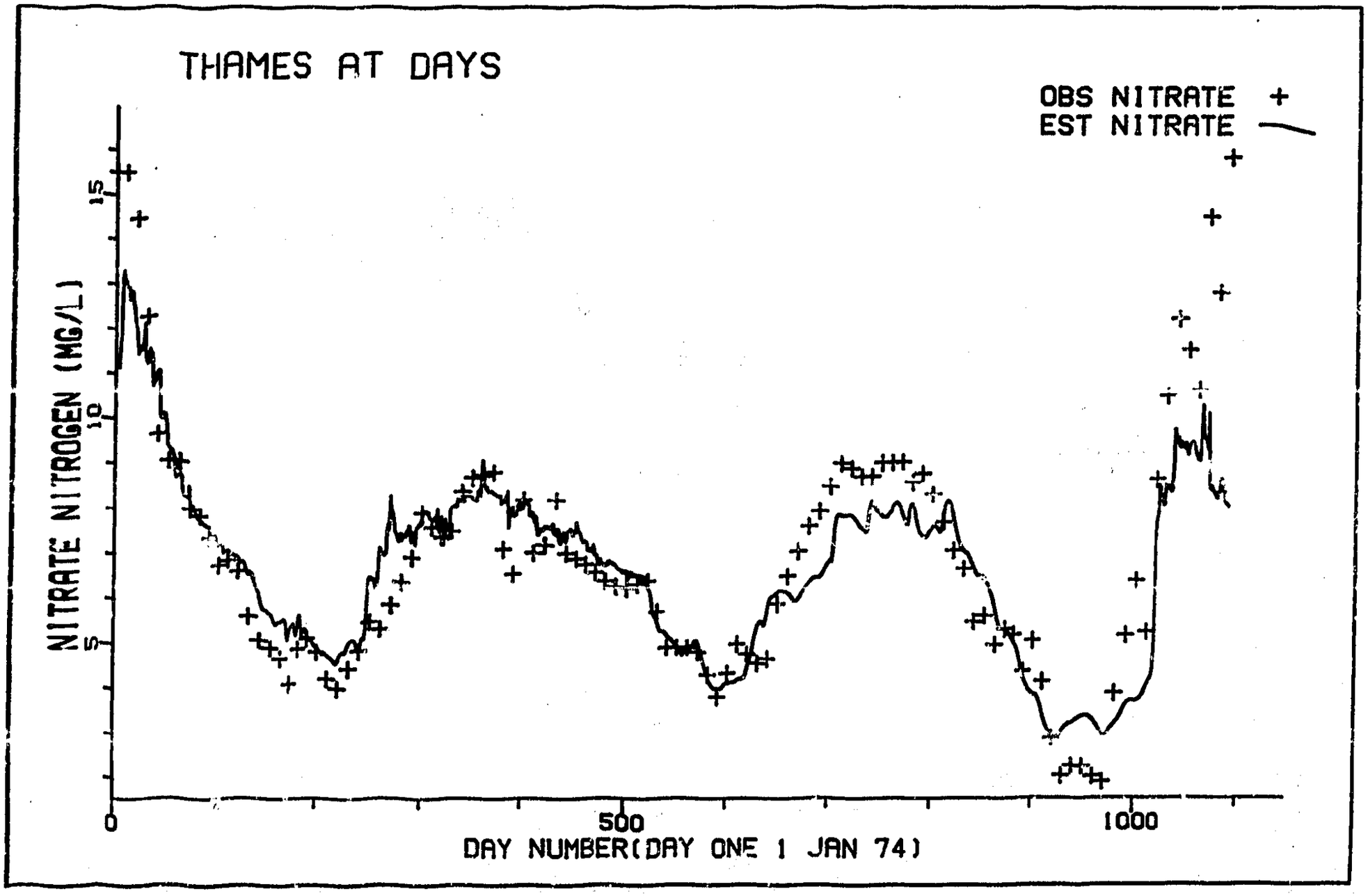


Figure 4.11 Simulated and Observed Nitrate (NO_3 as N) concentrations at Days Weir, 1974-1976 allowing for temperature dependent[3] denitrification processes.

column nitrogen losses. A modelling study has been undertaken (Whitehead and Hornberger, 1984) to develop a model of algal transport, growth and death in the Thames which may ultimately be incorporated into the Thames nitrate model.

The final Thames River basin model will incorporate surface runoff components and groundwater models. Given likely agricultural developments over the next twenty five years, simulations using the model will produce information on predicted nitrate concentrations at key locations along the river. Strategies for managing the nitrate problem can then be assessed using the simulation model.

TABLE 4.1 % Loss by denitrification within various reaches of the River Thames

	Year	Swinford	Caversham	Datchet	Teddington
Summer conditions	74	62	91	73	70
	75	67	92	73	83
	76	97	99	3	91
Winter conditions	74	5	13	13	12
	75	5	13	13	12
	76	37	75	53	40

4.3 Modelling algal dynamics in river systems

The objective of the Thames model described in Section 4.2.2 is to represent nitrogen sources and sinks as well as processes causing the redistribution of nitrogen in the river system. In rivers there is a natural cycle of nitrogen caused by the uptake during growth of algae and subsequent nitrogen release after algal death and decay; the dead algae are frequently flushed from the system or deposited on the bottom to be scoured and resuspended under high flow conditions.

Algal growth may also cause operational management problems in the Thames Water Authority since all of the abstracted water from the Thames is pumped into reservoir storage prior to distribution to water treatment plants. Algal growth in the reservoirs can affect water taste and smell and cause filtration problems in the treatment plants. The prediction of algal population dynamics is therefore of considerable importance in reservoir management.

In order to investigate algal growth in the Thames, a modelling study was undertaken during the workshop.

4.3.1 Background

Figure 4.12 shows the Thames and the tributary of the River Lea joining the main stream and the location of main sampling stations. The length of the main river is 236 kilometres with a fall of 108 metres. There are a considerable number of locks and weirs on the river. The weirs have a major effect on water quality since they determine the depth and hence the retention time within reaches. Except in times of high flow the depth regulation results in low water velocities and long retention times and hence ideal conditions for the growth of algae. The use of numerous small boats in the summer tends to increase the turbidity in the river and is likely to diminish algal growth.

Sewage and industrial trade effluent discharged to the Thames sytem above Teddington forms a high proportion of the total flow in periods of low flow. Whilst some water is abstracted for public water supplies from the River Thames above Oxford, from the River Kennet at Reading, and from a number of chalk and limestone springs, the largest abstractions are made from the reaches of the Thames between Windsor and Teddington. The main abstractor is the Metropolitan Water Division of the Thames Water Authority who have numerous intakes along these lower reaches.

4.3.2 Water quality and algal data

Data for the years 1974, 1975, 1976 have been obtained from the Thames Water Authority for water quality and algal levels at six sites along the Thames, namely Castle Eaton, Buscot, Swinford, Caversham, Staines and Teddington and these form the boundaries of five reaches, as indicated in Table 1. The data have been collected on approximately weekly basis by Thames Conservancy and chlorophyll a (by both methanol and acetone extraction), pheopigment, nitrate and silica data are also available. As indicated in Table 1 there are relatively low levels of chlorophyll a in the upper reaches of the Thames. However significant algal growth occurs between Swinford and Caversham and mean levels downstream at Caversham show no major increase. A strong seasonal pattern in algal growth occurs with major spring and autumn blooms of diatoms such as *Stephanodiscus*. In summer months green algae tend to dominate, and in 1976, which was a particularly severe year in hydrological terms with low flow rates, major blooms of *Microcystis* occur (Davis, 1977).

Nitrate concentrations also show significant seasonal variation with high concentrations in winter and low concentrations in summer months. During summer nitrogen is removed by bacterial action under anerobic conditions in the mud and by the uptake of nitrogen by macrophytes and algae. These processes depend on such factors as temperature which, for example, controls the rate of denitrification (Toms *et al*, 1975) and other variables such as flow rate and solar radiation which affect macrophyte and algal growth.

Nitrate in the Thames originates from point sources such as discharge of sewage effluent or from diffuse sources such as land drainage and rainfall. Owens *et al* (1972) have shown that over 80% of the nitrate in the Bedford Ouse Catchment (adjacent to the Thames Catchment and similar in being

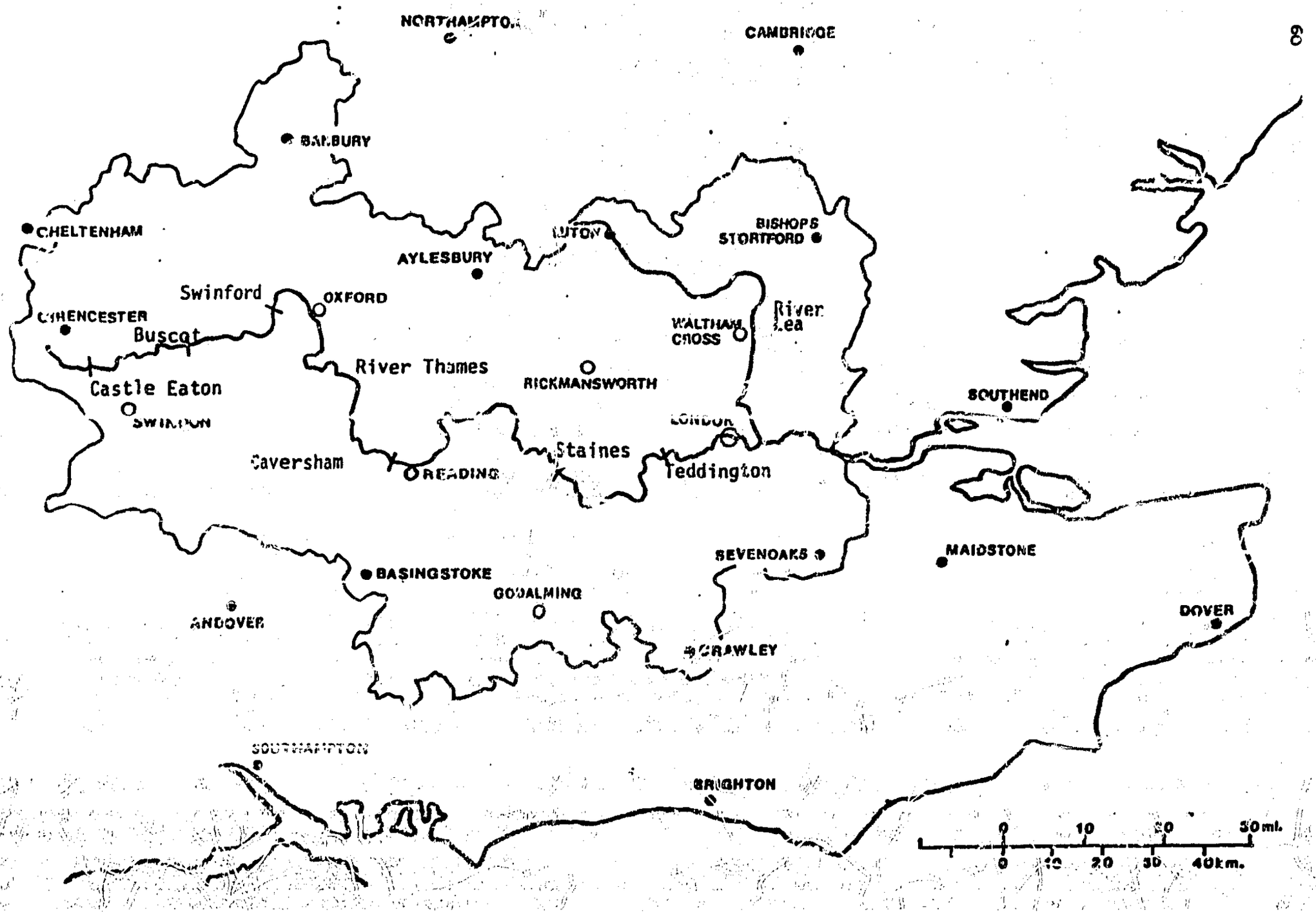


Figure 2.2 The Thames Catchment Area

largely agricultural) is derived from diffuse sources and Onstad and Blake (1980) have shown that 78% of the variance of nitrogen in the Thames is due to diffuse or non point sources.

TABLE 4.2 Weekly chlorophyll a (mg m^{-3}) concentrations for various reaches of the River Thames over the period 1974-76

	Reaches	Mean Value	Max Value
1 -	Castle Easton	10.6	60.2
2 -	Buscot	14.6	113.2
3 -	Swinford	29.1	284.6
4 -	Caversham	57.2	302.1
5 -	Staines	56.9	279.3
	Teddington	61.5	283.5

The nitrate concentrations in the River Thames are generally high and according to Lack (1971) are unlikely to limit algal growth. Lack also notes that diatom growth is unlikely to be limited by silica in the river. There are significant reductions in silica levels during the spring diatom blooms and in a reservoir situation the silica levels may fall to extremely low levels such that algal growth is limited. However in the river the silica levels remain at about 1 or 2 mg/1 and as such do not appear to constitute limiting levels.

One interesting feature of the Thames chlorophyll a data is the ratio of chlorophyll a obtained by acetone extraction to that of chlorophyll a obtained by methanol extraction. The ratio varies according to the species of algae since chlorophyll a by acetone extraction is a "relatively" more efficient extraction method for species such as blue-green or green algae compared with diatoms. The weekly chlorophyll a ratios have been numerically filtered. It was necessary to smooth the ratio using a low pass filter with a time constant of four weeks. The high frequency noise associated with measurement error is removed to some extent by the filtering process. The major fluctuations in the ratios indicate different species of algae occurring at different periods of the year. In general, the pattern of variation indicates diatom blooms in spring and autumn and summer blooms of greens in 1974 and 1975 and a bloom of blue-greens in summer 1976.

4.3.3 Application of modelling techniques

Algal distribution and growth processes in the River Thames have been the subject of research by a number of biologists this century (Fritsch 1902, 1903, 1905, Rice 1938, Kowalczewski and Lack, 1971, Lack 1971, Bowles, 1978). This research has not involved the use of mathematical models and conclusions have been based on correlation analysis of the data or mass balance approaches. Given the highly dynamic nature of the river and the non-linear processes governing algal transport and growth, it is necessary to use modelling techniques to obtain an adequate

description of the system. There have been few modelling studies of algal processes in rivers although the analysis of flow and quality data using modelling techniques has developed considerably in recent years (Thomann 1972, Beck and Young 1976, Whitehead *et al*, 1979, 1981). A useful approach is to establish a component mass balance across a reach of river:

$$\frac{dx(t)}{dt} = a\,u(t) - b\,x(t) \qquad (4.13)$$

where u(t) represents the upstream (input) algal concentration (mg/m^3); x(t) represents the downstream (output) algal concentration (mg/m^3); and a and b are parameters to be estimated. If algal growth and death processes were negligible (ie algae could be considered a conservative variable) then a and b would be related to the inverse of the time constant, T, for the reach ie $\frac{Q(t)}{V} = \frac{1}{T}$ where Q(t) is the flow rate in the reach and V is the reach volume.

However, as shown by previous studies, algal growth is significant in the Thames and one method of investigating this is to use the extended Kalman filter to estimate the a and b parameters. The EKF technique has been applied in a number of water quality studies (Beck and Young, 1976; Whitehead *et al*, 1981) and the approach has been discussed in detail in Section 2.2. The EKF is a recursive algorithm in which an estimate of the unknown parameter vector $\underline{\alpha}$ is undated while working serially through the data. The estimate $\hat{\underline{\alpha}}$ of $\underline{\alpha}$ at the kth instant in time is given by an algorithm of the following form:

$$\hat{\underline{\alpha}}_k = \hat{\underline{\alpha}}_{k-1} + G_{k|k-1}\{\underline{y}_k - \hat{\underline{y}}_{k|k-1}\} \qquad (4.14)$$

where the second term on the right hand side is correction factor based on the difference between the latest determinand measurement $\underline{y}_k$ and the estimate $\hat{\underline{y}}_{k|k-1}$ of that determinand derived from the model using estimated model coefficients obtained at the previous time point. $G_{k|k-1}$ is a weighting matrix whose elements are calculated essentially as a function of the levels of uncertainty (or error) specified for the model in the output response and the unmeasured input disturbances.

Applying the EKF to the algal data provided some useful information on the nature of the system. Figure 4.13 shows the estimated a and b parameters for the third reach and Figure 4.14 shows the observed and estimated algal concentrations.

The ratio b/a is a measure of the gain in the system so that when b is larger than a the gain is greater than unity and algal growth processes are significant. Conversely when the parameter b is less than the parameter a the gain is less than unity and death or sedimentation processes are dominant. In the case of reach five the parameters are similar until the spring of 1975 and then the b parameter diverged markedly from the a parameter suggesting major growth of algae during

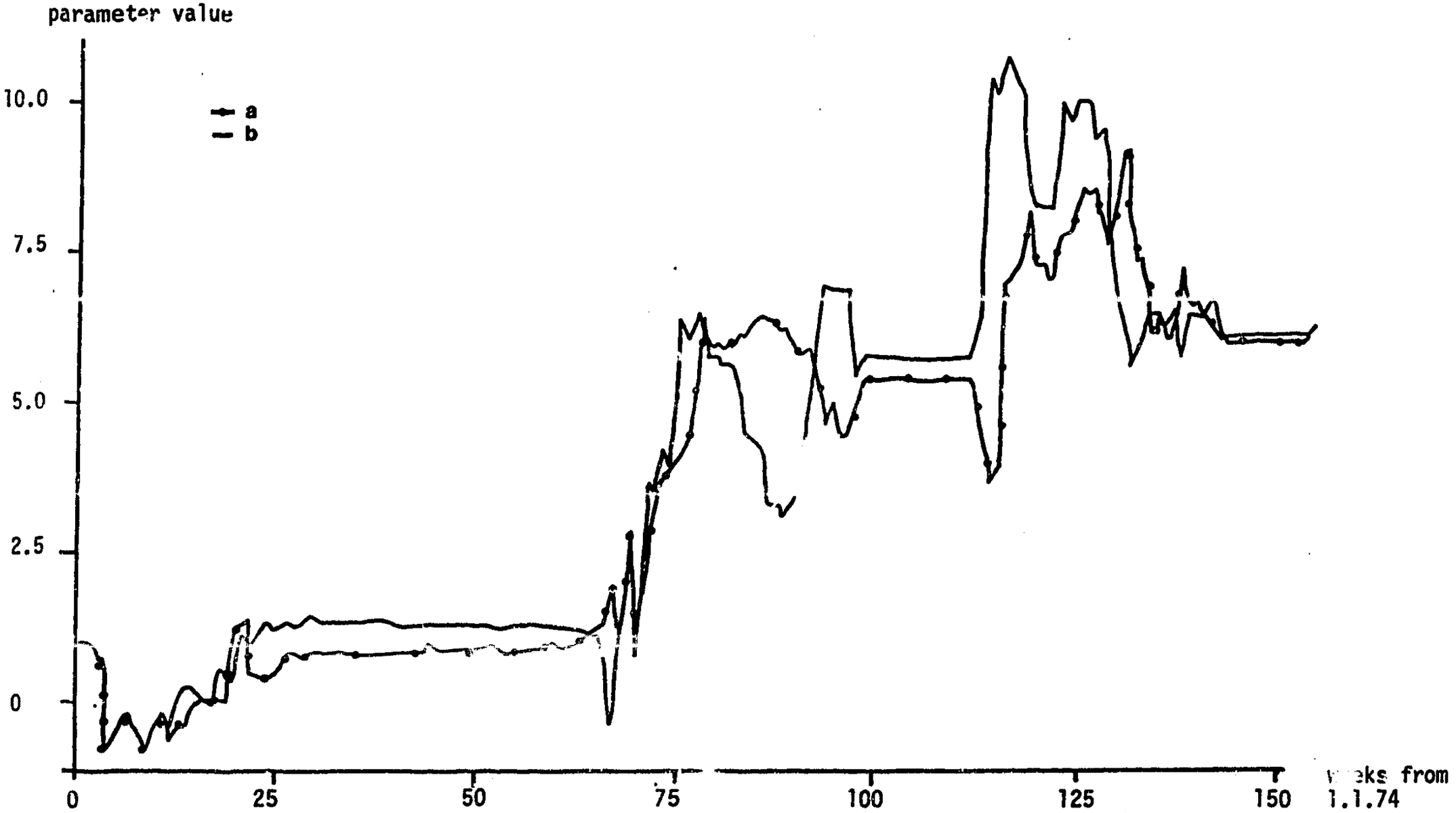


Figure 4.13 Parameters a and b from EKF estimation

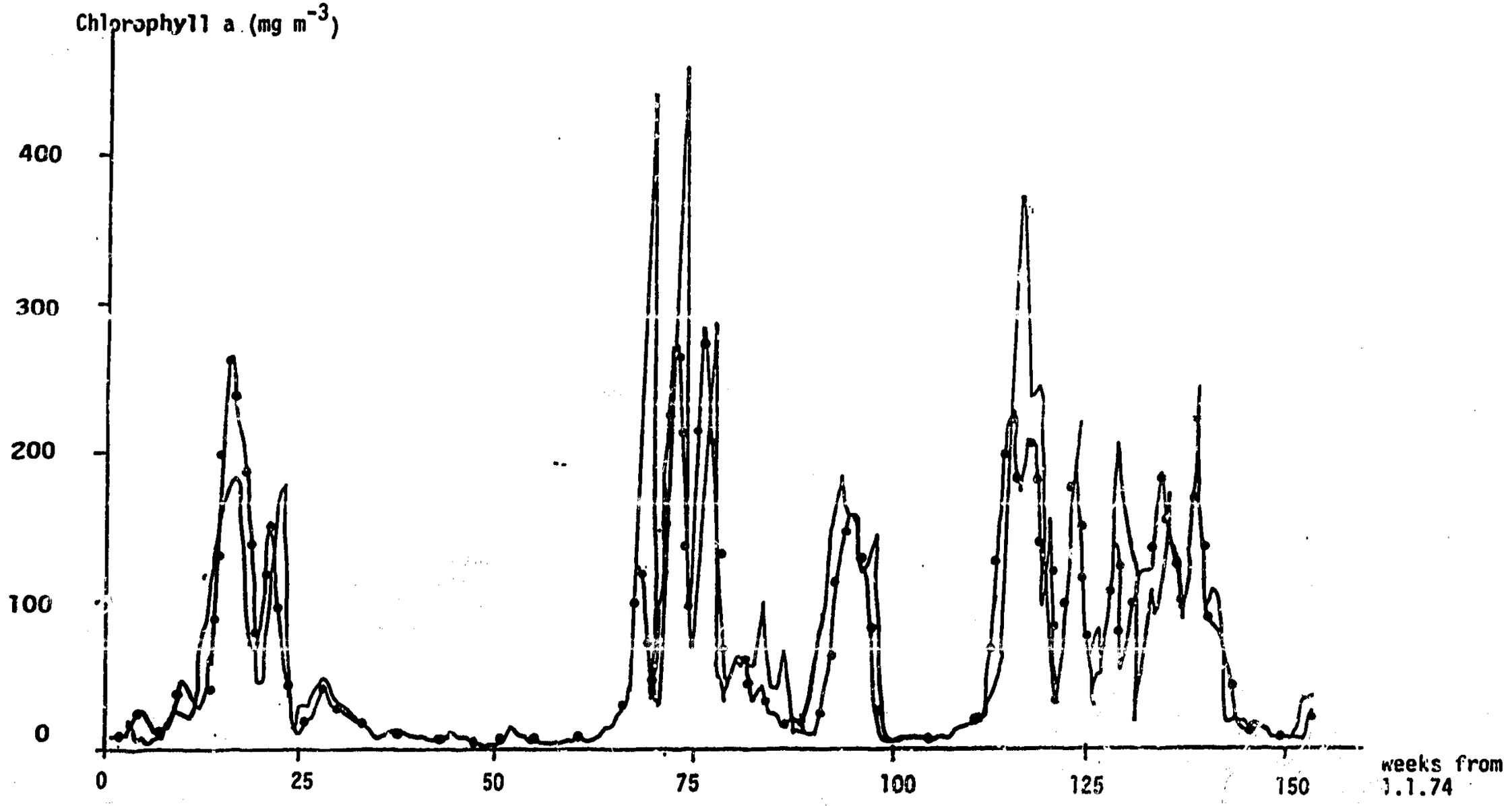


Figure 4.14 Estimated and Observed Chlorophyll a levels from EKF analysis applied to 5th Reach (Staines to Teddington)

this period. The high growth period in spring is followed by a loss of algae represented by the b parameter falling below the a parameter. This loss may be due to algal death or sedimentation. Spring diatoms are relatively heavy algae (containing silica in their cell structure) and in the lower reaches of the Thames, which tend to be less turbulent, the algae are likely to settle out on the bed of the river (Bowles, 1978).

It is clear from this analysis that complex processes of algal growth and algal removal are occurring along the river and in order to model such behaviour it is necessary to hypothesize mechanisms for these processes.

The extensive literature on processes controlling algal growth (Thomann, 1972; Bowles, 1978; and Lack, 1971) provide an essential background to this part of the research. Rather than take a standard model developed for a particular system the approach herein has been to evaluate the most likely factors controlling algal growth and losses and to represent these mathematically. The four factors considered particularly important for algal growth are:

(i) the growth coefficient;

(ii) the effect of solar radiation which under conditions of unlimited nutrients provides the main driving force for algal growth;

(iii) the effect of turbulence which tends to increase with increasing flow causing resuspension of sedimented material and reducing light penetration;

(iv) the self-shading factor in which an algal population grows to the point where light penetration is reduced by shading of the algae.

Initially it was assumed that nutrients such as nitrate, phosphorus and silica were not limiting in the Thames (Lack, 1971) and that silica was available in sufficient quantities to sustain algal growth.

The loss processes were assumed to be related to the concentration of algae via a first order decay term and mathematically the live and dead algae are represented as follows:

Live Algae

$$\frac{dx_1(t)}{dt} = k_1 Q_1(\ ) u(t) - k_1 Q_0(t)\, x(t) - k_2 x_1(t) + k_3 \frac{I(t)}{Q_0(t)} \left(\frac{k_4}{k_4 + (x_1(t))^{k_5}}\right) \left(\frac{I(t)}{k_6}\right)^{k_7} \exp\left(1 - \left(\frac{I(t)}{k_6}\right)^{k_7}\right) \qquad (6.14)$$

Dead Algae

$$\frac{dx_2(t)}{dt} = - k_1 Q_0(t) x_2(t) + k_2 x_1(t) - k_8 x_2(t) \qquad (4.16)$$

where $x_1(t)$ and $x_2(t)$ represent the live and dead algae respectively at the output (downstream) boundary of the reach, measured as chlorophyll a $\mu g \ell^{-1}$; $u_1(t)$ represents the input (downstream) algal concentration ($\mu g \ell^{-1}$); $Q_1(t)$ and $Q_0(t)$ represent the upstream and downstream flow rates; $I(t)$ is the solar radiation level (watts cm^2);

k_1 determines the residence time characteristics of the model such that $k_1 Q_0(t) = 1/\tau$ where τ is the residence time;

k_2 is the algal death rate;

k_3 is the growth coefficient;

k_4 is a half-saturation level for the self-shading function

$$\frac{k_4}{k_4 + (x_1(t))^{k_5}}$$

and k_5 is included as a power term on $x_1(t)$ to enhance the self-shading factor at high algal concentrations;

k_6 represents the optimal solar radiation level in the term;

$$(\frac{I(t)}{k_6})^{k_7} \exp(1-(\frac{I(t)}{k_6})^{k_7})$$

which accounts for the decrease in algal growth under low light intensity and the apparent decrease in growth under extremely high light intensity conditions in the Thames;

k_7 enhances the effect of this solar radiation term;

k_8 is included in the dead algae equation to account for the loss of algae by sedimentation.

An additional parameter k_9 is included in the model as a temperature threshold below which algal growth is zero.

Many researchers have developed phytoplankton growth models for simulation purposes. In general, the approach to parameter estimation has been to select parameters quoted in the literature and assume these values pertain to the system under investigation. Lederman *et al* (1978), however, applied non-linear parameter estimation techniques to data from batch cultures of phytoplankton to directly estimate model parameters. Such techniques are required to estimate unbiased parameters. An

alternative approach is to apply an instrumental variable algorithm to the differential equation model (Whitehead, 1981). Here the extended Kalman filter technique, applied above to the simple mass balance differential equation model given by equation (4.13), has been used to estimate model parameters. However, prior to the EKF analysis a sensitivity analysis has been conducted to investigate the effect of parameter uncertainty on the behaviour of the model and to analyse the uncertainty associated with model structure.

The generalized sensitivity analysis technique discussed in Section 3.2 can be utilised to identify the dominant parameters and hence mechanisms controlling the system behaviour. This is an important aspect of the modelling study since it is generally not possible to obtain reliable estimates of the large number of parameters in most simulation models. It is preferable to fix the values of those parameters which are thought to be well known and to then optimize the remaining parameters. Up to now there has been no systematic method of selecting the subset of parameters for optimization. A trial and error procedure is normally used to select these parameters but given the non-linear nature of most simulation models such an approach can present problems of interpretation and is certainly not rigorous. The generalized sensitivity analysis can aid in this parameter selection to ensure that the optimal set of parameters is obtained.

4.3.4 Application of sensitivity analysis to the Thames algal model

In applying the sensitivity analysis approach to the Thames algal modelling study, it is first necessary to define the system behaviour. The two important features of algal growth within the river is the presence of a spring bloom and the subsequent fall to relatively low levels after spring, in early summer. Simulations are classified as a behaviour if the algal concentration, x_1 is at any time, above 100 $\mu g/\ell$ and below 400 $\mu g/\ell$ during a 5 week period in spring and if, in addition, x_1 falls below 100 $\mu g/\ell$ and remains below this level for at least two weeks during the five weeks after the spring bloom.

The model parameters were selected initially on the basis of published information such as travel times for the Thames determined by the Water Authority or growth rates for algae in the Thames. As previously discussed there is considerable uncertainty associated with many of the parameters in the model. In the case of growth rates, for example, Swale (1962), measured a growth rate for *Stephanodiscus Hantzschie* of 0.46 $days^{-1}$ and Bowles (1978)determined a growth rate for *Asterionella* of 1.28 $days^{-1}$ from studies on the Thames. Lund (1949) also determined a growth rate for *Asterionella Formosa* of 1.73 $days^{-1}$ under field conditions. The situation in the Thames is complicated by the changing nature of the river with relatively slow flow in the lower reaches compared with the flow in upper reaches between Buscot and Swinford. Selecting a value for the Thames is therefore particularly difficult and as an *a priori* estimate the value of 1.43 $days^{-1}$ has been selected.

A complete list of parameter values for the Monte Carlo simulation runs is given in Table 4.3. In the table the mean parameter values are given. However, in the Monte Carlo runs the parameter values are selected randomly assuming a rectangular distribution with a variance of ± 50% of the mean of the parameter. This ensures that a wide spread of parameter values is selected and that behavioural patterns are fully explored.

TABLE 4.3

Monte Carlo simulation results for reach 5 (Staines to Teddington); (statistics on maximum distribution separation invalid in run 4)

Monte Carlo Simulation Runs		1		2		3		4
Critical $d_{m,n}$ at 90% confidence level		0.326		0.430		0.470		invalid $d_{m,n}$
Parameter Value (P) Distribution Separation (S)		P	S	P	S	P	S	P
k_1	related to travel time τ, $k_1 Q_0 = 1/\tau$	0.5	0.2	0.16	0.14	0.16	0.34	0.16
k_2	algal death rate (weeks $^{-1}$)	0.3	0.2	0.3	0.16	0.3	0.19	0.6
k_3	algal growth rate (weeks^{-1})	10	0.15	8.0	0.33	10	0.48	12
k_4	algal saturation level ($\mu g \ell^{-1}$)	100	0.17	100	0.1	100	0.17	100
k_5	power in saturation term	2	0.73	2.5	0.68	3	0.87	4
k_6	optimal solar radiation (watts cm^2)	20,000	0.51	13,000	0.5	15,000	0.51	10,000
k_7	power in light attenuation term	2	0.30	2	0.42	2	0.31	2
k_8	sedimentation rate (weeks^{-1})	0.3	0.13	0.3	0.26	3	0.25	0.3
k_9	temperature threshold effect (^{0}C)	8	0.11	8	0.16	8	0.17	8
% behaviour (based on 100 simulations)		48%		79%		86%		98%

Table 4.3 shows the parameter values used in four Monte Carlo simulations together with the maximum separation between the parameter distributions and the critical separation $d_{m,n}$ at the 90% confidence level. It is particularly interesting to note that relatively few parameters appear to be significant in determining behaviour. Over the four simulations only three parameters are clearly identified as critical, these being the growth rate k_3, the power term in the saturation factor k_5, and the optimal solar radiation levels k_6. In the first simulation only 48% of the runs satisfy the behaviour criterion and from analysing the means of the behaviour producing parameters it is possible to determine whether to increase or decrease the parameters in order to increase the percentage of behaviours. For example, in the case of the power term parameter in the saturation function, the standardized mean under the behaviour is 0.54 as shown in Table 4.4 suggesting that this parameter should be increased. By increasing this parameter the shape of the saturation function is altered thus enhancing the effect of the saturation level. Similarly in the case of k_6, the optimal solar radiation level, the mean under behaviour is - 0.56 suggesting a reduction in this parameter. The Monte Carlo simulations therefore, can be used as a crude estimation procedure and the % of behaviours increased from 48% to 98% over the four runs using this approach.

TABLE 4.4 Statistics for 9 parameters in simulation run 1

Parameter	Standardized Mean under Behaviour	Standardized Mean under Non-Behaviour
k_1	.21	- .10
k_2	- .24	.57
k_3	- .10	.18
k_4	.62	.21
k_5	.54	- .70
k_6	- 56	.42
k_7	.34	- .24
k_8	- .84	.53
k_9	- .11	- .18

From a systems analysis point of view what is particularly significant is that only three of the nine parameters control system behaviour. In most modelling studies of ecological or hydrological systems it is conventional to give equal weight to model parameters. Moreover, in many simulation studies, a trial and error procedure of model calibration occurs in which a subset of the parameters are adjusted until a reasonable model fit is obtained. With large complex models this process can be particularly difficult because of interactions between parameters and mechanisms. The generalized sensitivity analsis approach can

therefore be used in this situation to determine in a systematic manner the dominant parameters controlling behaviour.

The final stage of any parameter estimation study should be the use of a statistically based technique to determine an unbiased optimal set of parameters. Unfortunately in the presence of measurement noise and system noise any parameter estimation algorithm is limited in application. In the case of the Thames algal model it was impossible to apply a technique such as the extended Kalman filter to estimate all nine parameters. The EKF technique applied in this situation gave parameter values which were either clearly incorrect or showed colinearity in which one parameter increased as another decreased to cancel out its effect. Such parameter behaviour is indicative of an unobservable system and is a common feature of most simulation models. The question of observability is important here since as indicated by the simulation study only three parameters are significant in controlling behaviour in the Thames model.

Thus in order to obtain reasonable parameter estimates the EKF must be applied to these three critical parameters with the remaining parameters set to the best values obtained from either the Monte Carlo simulation analysis or from laboratory or field measurement.

4.3.5 EKF applied to reduced model

The estimation results obtained by the EKF for the fourth and fifth reaches are shown here. Figures 4.15, 4.16 and 4.17 for the fifth reach, show respectively the estimated and observed state variable, x_1, the estimated 'dead' algal state, x_2, and the parameter estimates obtained from the EKF analysis. In general, the state estimate corresponds well with the observed chlorophyll a values and the parameters k_3, k_5 and k_8 are reasonably time invariant. The parameters show some movement at week 90 and this corresponds with a data period when the model estimate is below the observed levels; in this situation the parameters are adjusted by the EKF algorithm to compensate for the lack of fit. Similarly in the fourth reach the state estimate, x_1, compares reasonably with the observed values and the dead algal state variable agrees to some extent with the pheopigment levels, shown in Figure 4.18. The pattern of behaviour and concentration levels are similar and it may be possible to use the pheopigment as a surrogate measure of dead algae within the reach.

The simulation results and parameters obtained by the EKF analysis for the fourth reach, as shown in Figures 4.19 and 4.20 are more variable than those for the fifth reach. The power in the algal saturation term reduces from 4 down to 3.3 and the growth coefficient increases over the 1976 summer period. These changes may be due to the different types of algae dominating the river system. For example in summer 1976 there was a major bloom of *Microcystis* and self-shading is different because of the different size, clustering and buoyancy characteristics of *Microcystis* compared with other algae.

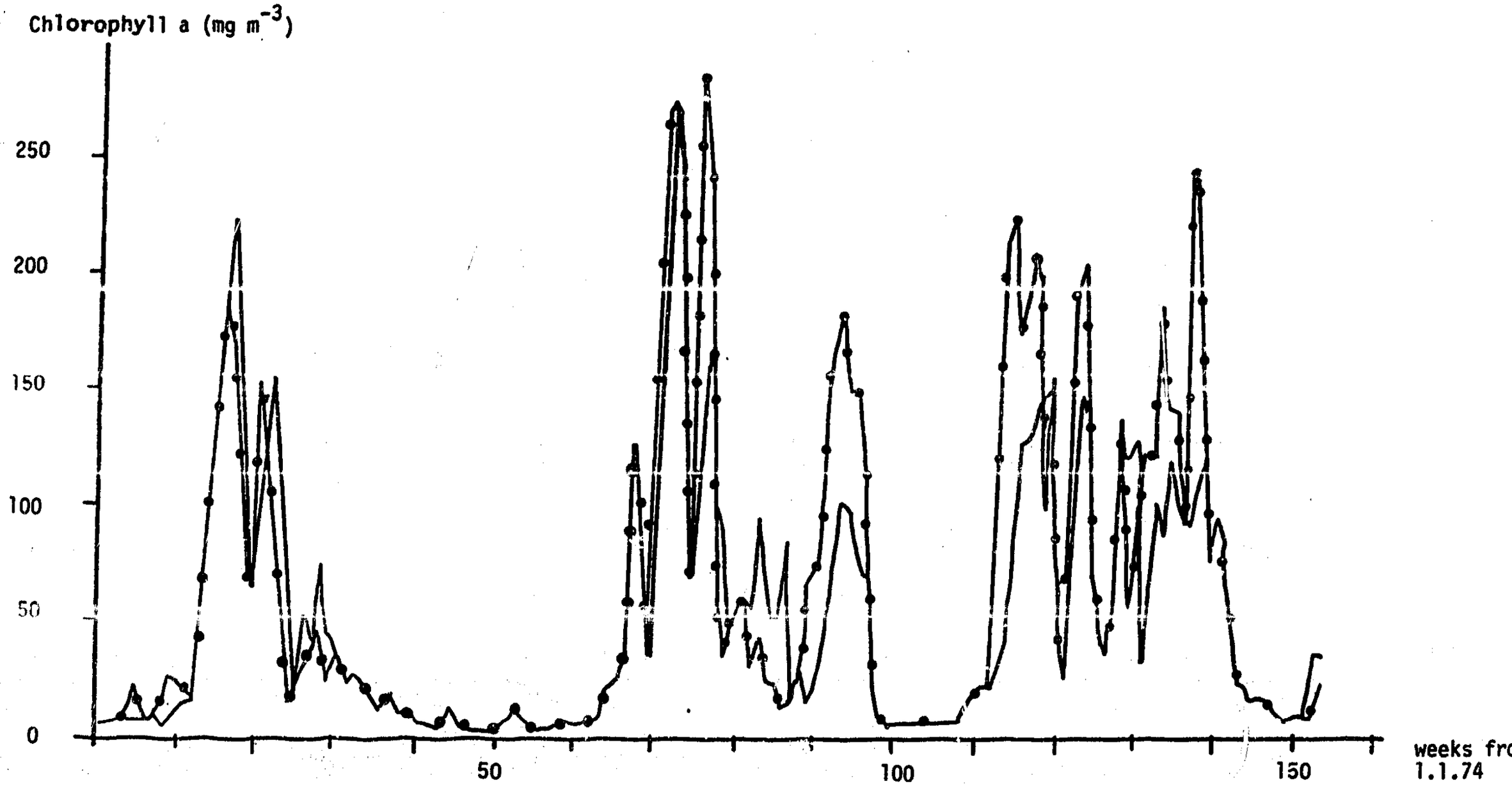


Figure 4.15 Estimated and Observed Chlorophyll a for 5th Reach

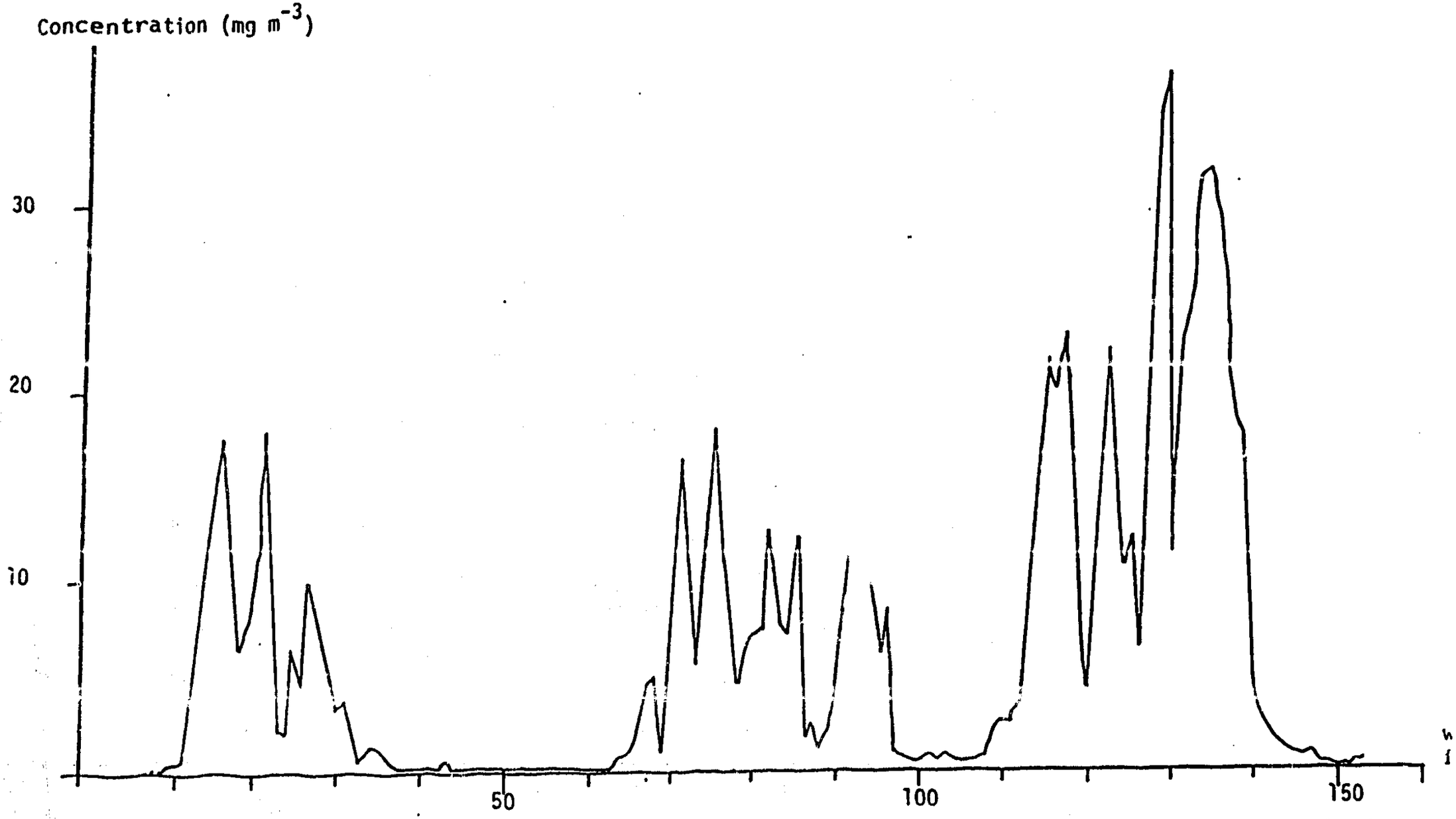


Figure 4.16 Estimated state x_2 (dead algae) in 5th Reach.

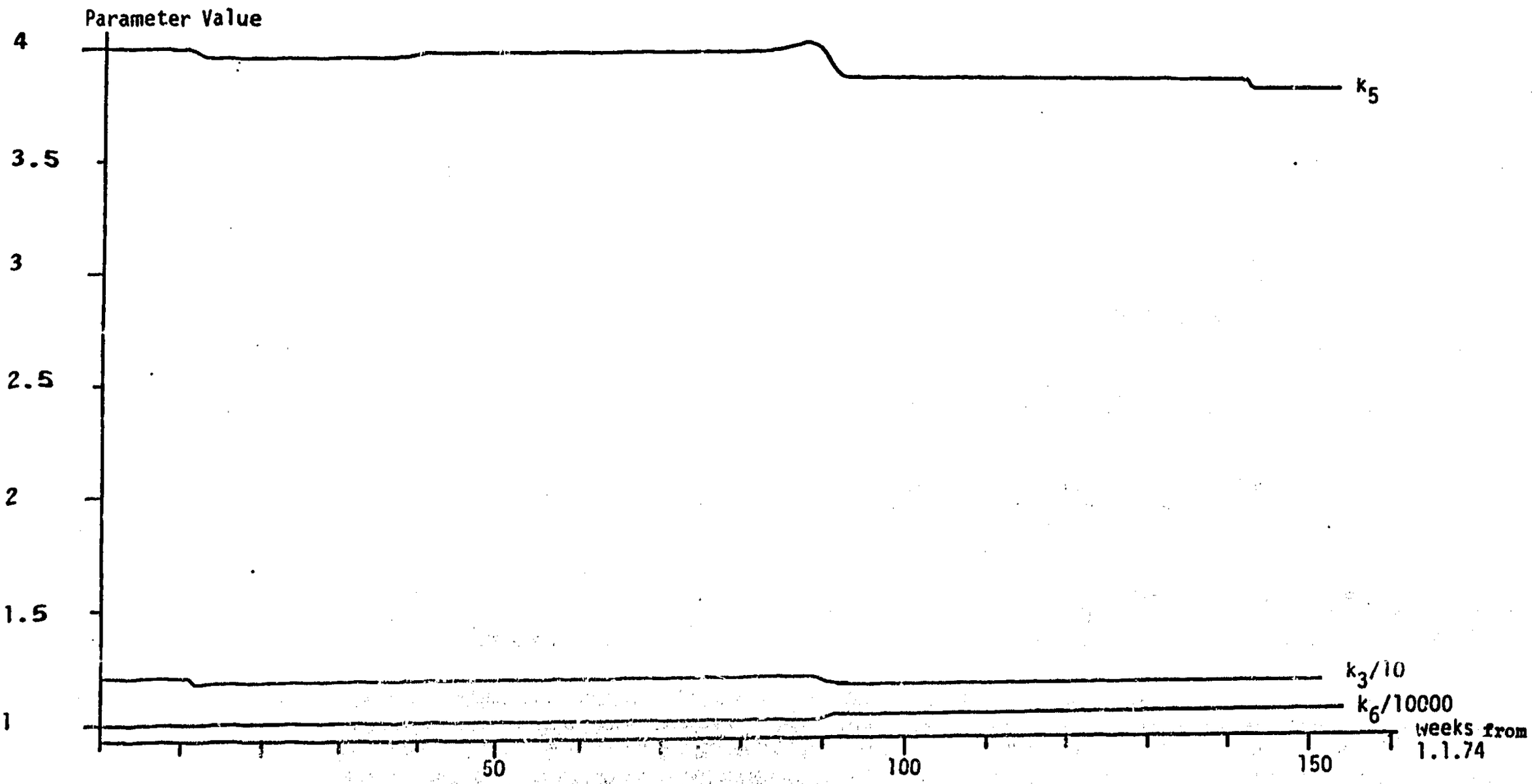


Figure 4.17 Estimated model parameters k_3, k_5 and k_6 for 6th Reach

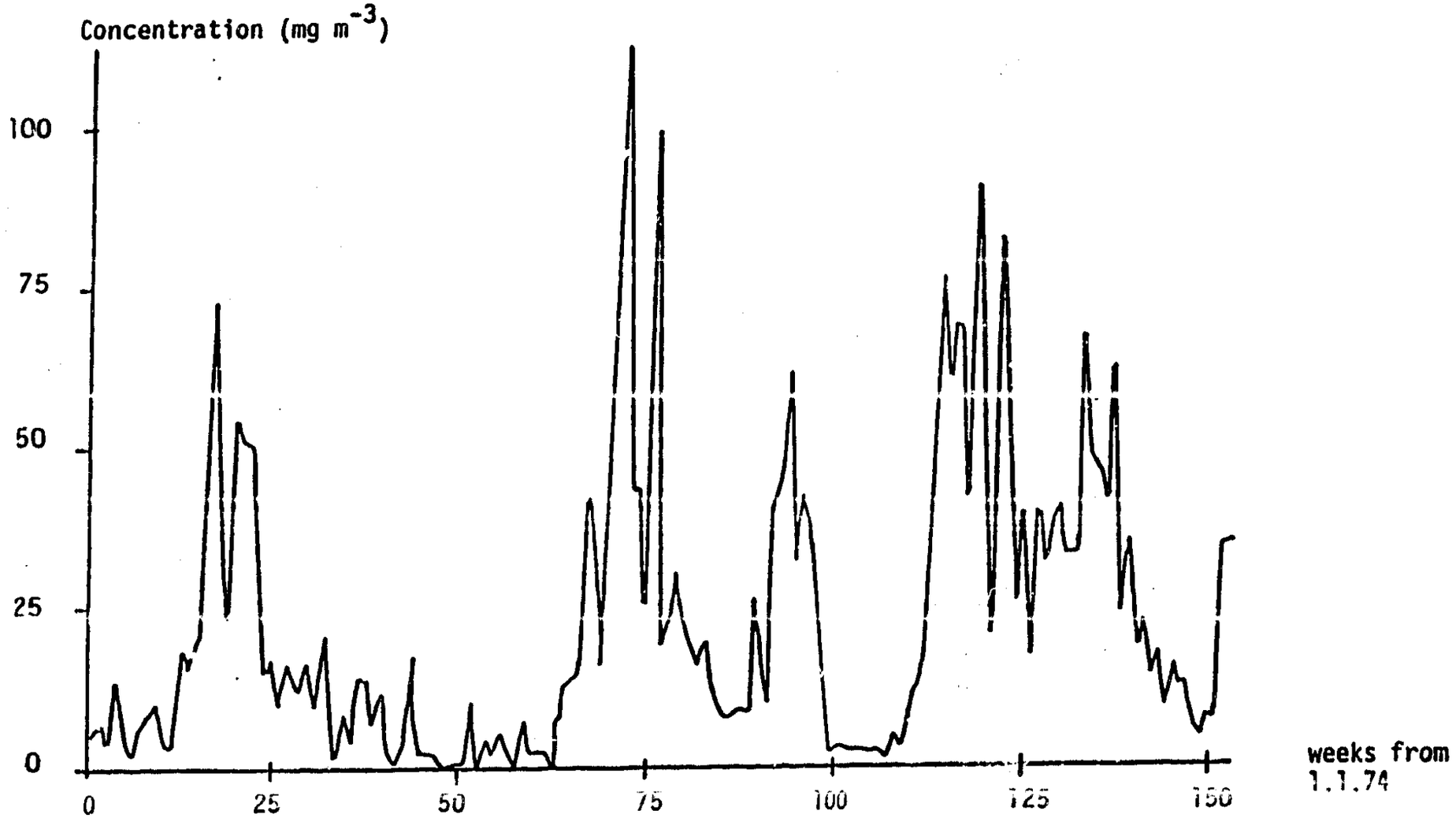


Figure 4.18 Pheopigment concentrations in River Thames over 1974, 75 and 76

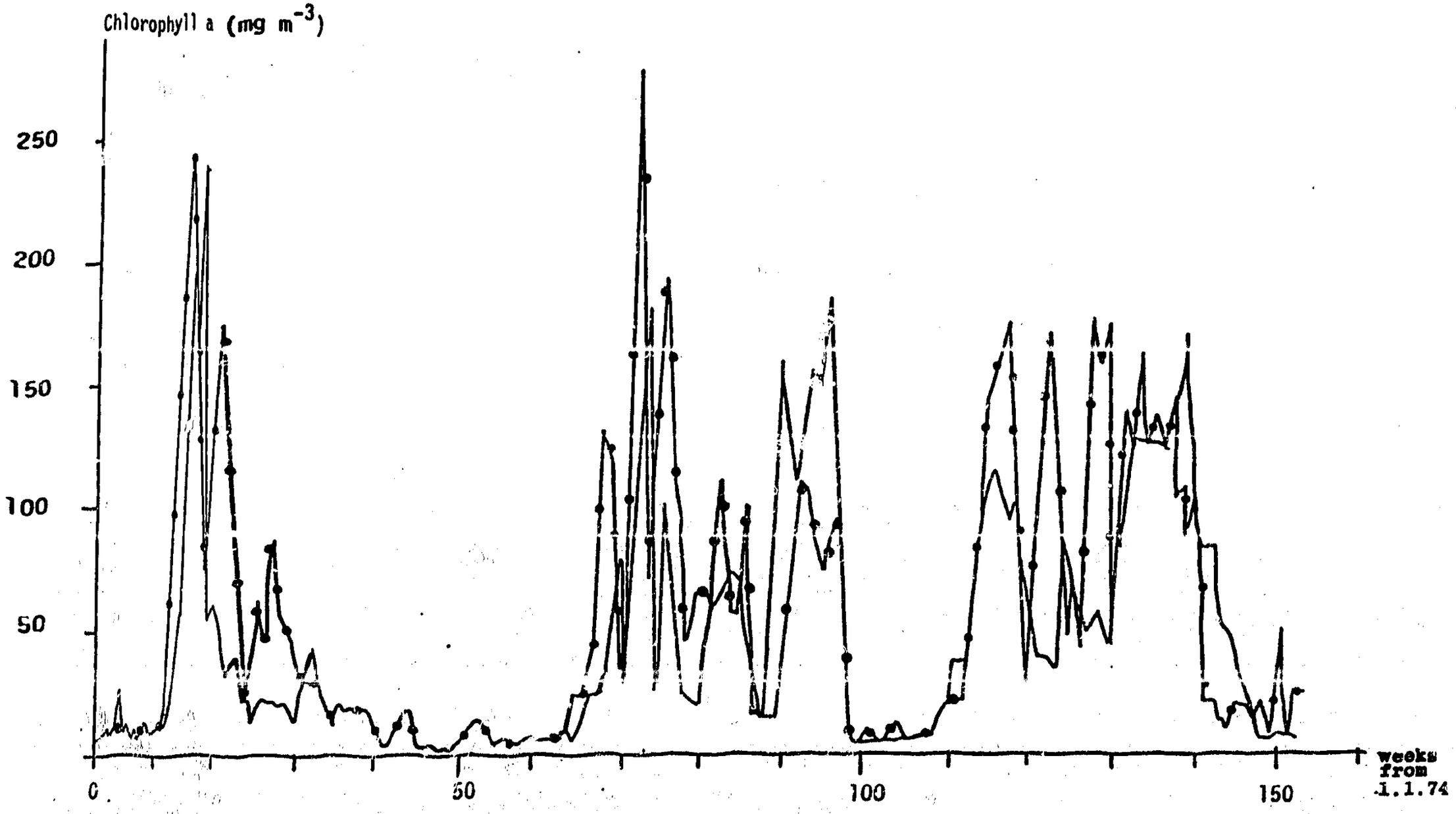


Figure 4.10 Estimated and Observed Chlorophyll a levels in 4th Reach

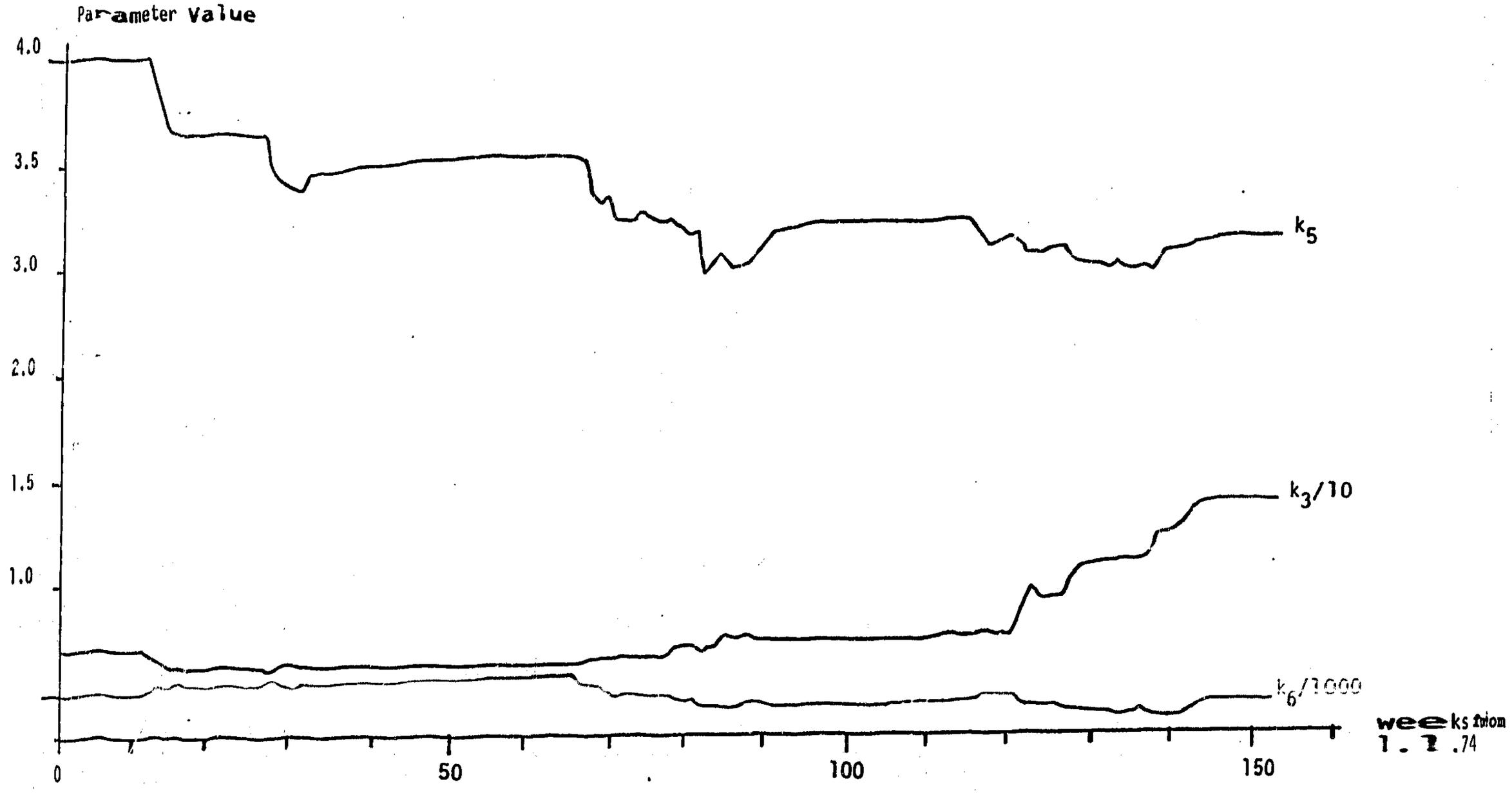


Figure 4.20 Estimated model Parameters K_3, k_5 and k_6 for 4th Reach

4.3.6 Conclusions

The complex dynamic behaviour of algae within river systems has been studied using a number of systems analysis techniques. Where transportation is a dominant factor controlling system behaviour, time series approaches are suitable. However, where algal growth processes dominate a mechanistic model is required to account for the highly non-linear behaviour. In this situation model identification and estimation is particularly difficult and a generalized sensitivity analysis technique can be used to determine the important parameters and hence restrict the number of parameters requiring estimation.

In the case of the Thames algal model three significant parameters have been identified out of the nine model parameters using the generalized sensitivity analysis technique. Having identified these parameters the EKF technique was applied to estimate final parameter values.

The application of the generalized sensitivity analysis approach prior to EKF analysis is suggested as a valuable approach, providing information on parameter uncertainty which can be used to reduce the estimation problem to a manageable level.

4.4 A generalised sensitivity analysis of the activated sludge treatment plant

4.4.1 Introduction

The application of the generalised sensitivity analysis to algal dynamics illustrated the advantages of the approach in determining the dominant parameters controlling system behaviour. The technique is particularly suited to these 'ill-defined' biological systems where processes and parameters are poorly understood.

The activated sludge treatment is another system controlled by biological processes and in this section of the proceedings we consider the application of the sensitivity analysis to this process.

The model considered includes components for simulating nitrification dynamics and dissolved oxygen dynamics in a well-mixed aeration tank and for accounting for a dynamic balance in a settling tank. A range of model parameters generally representative of the Norwich Sewage Treatment Plant (STP) (England) was selected. In addition, the behaviour of the Norwich plant was characterized by estimates of the range of variation of the system state variables under nearly constant sludge wastage rate, recycle rate and blower volumetric flow rate and the open loop consistency of the model was evaluated.

4.4.2 The mathematical model

Nine state variables are modelled. The differential equations of the model are given in Table 4.5. The equations for nitrogen dynamics in

TABLE 4.5. Equations of the Simulation Model

State Variables	Equation
1. Aeration Tank	
x_1, NH_3 - N	$\dot{x}_1 = \beta(u_1 - x_1) - \mu_1 x_3/Y_1$
x_2, NO_3 - N	$\dot{x}_2 = \beta x_2 + \mu_1 x_3/Y_1$
x_3, Nitrifiers	$\dot{x}_3 = \beta r x_9 - \beta(1 + r)x_3 + \mu_1 x_3 - k_4 x_3$
x_4, BOD	$\dot{x}_4 = \mu_2 x_5/Y_2 + \beta(u_2 - x_4)$
x_5, MLSS	$\dot{x}_5 = \mu_2 x_5 - b x_5 - \beta(1 + r)x_5 + \beta r x_7$
x_6, DO	$\dot{x}_6 = k_a(O_s - x_6)u_a - \alpha\mu_2 x_5 - \delta_2 x_5 - \beta(1+r)x_6$
2. Settler	
x_7, MLSS	$\dot{x}_7 = [1 - (\frac{x_7}{x_m})^3] \frac{Q_r + Q_w}{x_8} (C_r^{MLSS} x_5 - x_7)$
x_8, Sludge volume	$\dot{x}_8 = (\frac{x_7}{x_m})^3 \frac{Q_r + Q_w}{x_7} (C_r^{MLSS} x_5 - x_7)$
x_9, Nitrifiers	$\dot{x}_9 = (r + w) \frac{Q_I}{x_c} (C_1 x_3 - x_9)$

$$\mu_1 = \hat{\mu} \left(\frac{x_1}{k_{m_1} + x_1}\right)\left(\frac{x_6}{k_{m_6} + x_6}\right)$$

$$\mu_2 = \left(\mu_m \frac{x_6}{\theta + x_6}\right)\left(\frac{x_4}{k_{m_4} + x_4}\right)$$

$$C_r^{MLSS} = \frac{r + P_m + w\,(1 - P_m)}{r + w}$$

$$C_1 = \frac{r + P + w\,(1 - P)}{r + w}$$

$$P_m = 1 - 1.6\times10^{-5}\,(Q_I - Q_w)$$

the aerator follow those developed by Beck (1978) but the number of compartments is reduced from five to three. The equations for BOD and MLSS are in a commonly used form (eg see Andrews, 1974; Stensrom and Andrews, 1979). Dissolved oxygen is modelled following Olsson and Andrews (1978) and Tyteca *et al* (1977). The only "non-standard" components of the entire model are those representing the clarifier. The simplest approach to modelling the final settler strategies involving the recylce rate cannot be examined even cursorily with such a simplification. A theory presented by Kynch (1952) can be applied to model the thickening process in the settling tank (eg Busby and Andrews, 1975) but this requires treating the process as vertically distributed and thereby necessitates a rather large number of compartments to be modelled. We opted to utilize a modification of a CSTR suggested by Hornberger and Spear (1980) for representing a sediment compartment in a simplified simulation model of an estuarine ecosystem. This formulation divides the mass balance in the settler between sludge volume changes and concentration changes. At low concentrations, additions of solids are predominantly apportioned to increasing concentration in the sludge blanket but as a maximum concentration is approached additional inputs serve to increase the volume of sludge while the concentration remains almost constant. This simple formulation retains some of the important features of the more complex models and allows investigation of control strategies for the activated sludge process.

The parameters of the model are listed in Table 4.6. The limits of the rectangular probability distributions of parameters used in the Monte-Carlo simulations are also listed as are references which suggest these bounds for the parameters. The value of X_m, the maximum allowable concentration in the settler, was taken to be representative of the Norwich treatment plant. Values for recycle rate, wastage rate and air flow rate are likewise consistent with those for Norwich.

Constants and input concentrations and flow rates for the model are listed in Table 4.7. These are again consistent with observed conditions at the Norwich plant as presented in Beck *et al* (1978).

4.4.3 Results

Five hundred Monte-Carlo replications were run. Each simulation was for a period of six days. Initial condition transients were allowed to decay for one day and then the states were observed over a five day period and outcomes classified as either B or $\bar{B}$ according to whether the computed state variables stayed within the bounds given in Table 4.8 or whether one or more of these conditions was violated. Members of the behaviour class B are then open-loop consistent. That is, the trajectories stay within reasonable bounds. In most cases, these bounds are broadly consistent with observed conditions at the Norwich plant under conditions of little variation in the control variables, specifically those representing nitrifier concentrations and

TABLE 4.6 Parameters

Parameter	Range of Values	Units	Reference
1. Process parameters			
Y_1 (yield coefficient)	.006 - .06	--	1
$\hat{\mu}$ (nitrifier specific growth rate)	.02 - .08	h^{-1}	1, 2
k_{m_1} (half-rate coefficient for NH_3)	1.2 - 10	gm^{-3}	3
k_4 (rate coefficient for BOD)	.002 - .015	h^{-1}	1
μ_m (MLSS specific growth rate)	.1 - .4	h^{-1}	4, 5, 6
b (MLSS death rate)	.002 - .01	h^{-1}	4, 5, 6
k_{m_4} (half-rate coefficient)	25 - 200	gm^{-3}	4, 5, 6
k_a (BOD coefficient)	1.25×10^{-4}-7.5×10^{-4}	--	5
α (yield coefficient)	0.02 - .30	--	5
δ Yield coefficient	.50 - 1.42	--	5
p (Clarifier efficiency)	.88 - .95	--	1, 7
θ (rate coefficient DO/BOD)	.25 - 1.0	gm^{-3}	3
k_{m_6} (half-rate coefficient DO/NH_3)	.25 - 2.0	gm^{-3}	2, 3
y_2 (Yield coefficient)	.4 - .7	--	4
x_m (maximum MLSS concentration)	7000 - 10,000	gm^{-3}	7
2. "Control" variables			
r (recycle ratio)	.7 - 1.1	--	7
w (wastage ratio)	0 - .03	--	7
u_a (air input)	3×10^{5}-4×10^{5}	m^3d^{-1}	7

1. Poduska and Andrews (1975)
2. Stenstrom and Andrews (1979)
3. Jorgensen (1979)
4. Olsson and Andrews (1978)
5. Tyteca *et al* (1977)
6. Sincic and Bailey (1978)
7. Beck (1978)

TABLE 4.7 Constants and Inputs

1. Constant

V_A, aeration tank volume	8320 m^3
V_3, settler volume	4000 m^3
O_s, oxygen saturation	8.65 gm^{-3}
δ_1, stoichiometric coefficient	4.5 gO_2/gNH_3 - N

2. Inputs

	Mean	Diurnal Amplitude
Q_I, sludge inflow rate	2×10^{4} m^3d^{-1}	2×10^{4} m^3d^{-1}
u_1, NH_3 inflow conc.	40 gm^{-3}	10 gm^{-3}
u_2, BOD inflow conc.	300 gm^{-3}	150 gm^{-3}
u_3, DO inflow conc.	1 gm^{-3}	--

TABLE 4.8 Behaviour Criteria - Bounds on State Variables

State Variable	Bounds
x_1	0 - 40
x_2	5 - 50
x_3	0 - 2000
x_4	0 - 40
x_5	1500 - 4500
x_6	0 - 8
x_7	3000 - 10000
x_8	400 - 3600
x_9	0 - 2000

DO levels, we had little or no data from the Norwich plant and these variables were either unconstrained or very wide bounds were set in order to reject only their most extreme excursions. The behaviour conditions are thus chosen to isolate parameters that are known to be critical for open loop consistency. Of the 500 replications, 354 satisfied the behavioural conditions and 146 were classified as non-behaviours.

The statistical analysis of the results followed the procedure given in Spear and Hornberger (1980) and Hornberger and Spear (1981) and summarized in Section 3.2. The off-diagonal elements of the covariance matrix of parameters which yielded simulations in the B category were uniformly small and so information on parameter separation under the behavioural classification can be taken directly from the univariate statistics on individual parameters.

Taking the 99% confidence level of the two sample Kolmogorov-Smirnov statistic ($d_{m,n}$ = .16, where m and n are the sample sizes of the two groups) as the criterion for judging importance, four parameters stand out (see Table 4.9). Two of these, k_a and Y_2, are process parameters and two, r and w, are control parameters. The largest separation by far is in the oxygen transfer coefficient, k_a, and the next most important parameter is the MLSS yield coefficient, Y_2.

Of the 146 non-behaviours, 59 were caused by exceeding the allowable bound on the volume of sludge in the settler, 54 by BOD, 17 by dissolved oxygen, 11 by NO_3-N and 5 by MLSS in the aerator.

4.4.4 Discussion

The first implication of the results of the generalized sensitivity analysis is that the critical uncertainties in terms of simulating the activated sludge process with the present model structure reside in the DO-BOD-MLSS dynamics. This is the interpretation given to the generalized sensitivity results in previous studies (Spear and Hornberger, 1980) with the idea that further data collection/research efforts can be assigned priorities by using the sensitivity rankings.

To some extent the conclusion that the DO-BOD-MLSS dynamics are critical may be conditioned by the 5 day duration of the simulation. The choice of five days for the open loop simulation was dictated by the nature of the Norwich data in that this was the longest period over which the control variables, r, w and u_a were maintained at reasonably constant levels on a number of different occasions. It could be that some of the nitrification process parameter bounds would have been narrowed had data been available to allow simulation over longer periods. However, it is likely that any such errors will lead to conservative results in the eventual estimate of the closed loop behavioural probability, ie the probability of adequate closed loop behaviour is most likely to increase with a narrowing of the bounds defining open-loop consistency.

TABLE 4.9. Kolmogorov-Smirnov Statistic for Individual Parameters

Parameter	$d_{2,n}$
k_a	0.36
Y_2	0.27
r	0.22
w	0.21
b	0.16
μ_m	0.16
θ	0.15
k_{m_1}	0.13
Y_1	0.12
δ	0.12
α	0.11
x_m	0.11
k_{m_6}	0.10
p	0.10
u_a	0.09
k_{m_4}	0.08
k_4	0.08
$\hat{\mu}$	0.07

The second implication of the results is that control action may be expected to improve the response of the activated sludge process as judged by the model presented here. The overriding importance of the oxygen transfer coefficient reinforces the need for design of dissolved oxygen controllers (Marsili-Libelli, 1980). The importance of wastage rate in the sensitivity analysis is somewhat specious in that this undoubtedly arises because extreme low values of w lead to excess sludge volume in the clarifier, a condition that can easily be remedied by resetting the wastage on a daily basis. The fact that the recycle rate was important, however, points to the need for control of this variable in regulating the DO-BOD-MLSS dynamics of the aerator.

Finally, the sensitivity results may be interpreted as a crude form of parameter estimation in which the distributions of k_a and Y_2 should be altered to be consistent with the uncontrolled behaviour of the system prior to addressing the issue of control system design. That is, only those process parameter sets which are open-loop consistent can be examined in an expanded analysis in which parameters describing controllers in the closed loop are included. Thus, in designing controllers for poorly-defined systems in which uncertainties in process parameters are likely to be critically important, the preliminary sensitivity analysis described here is a necessary precursor to the full design problem.

5. CONTROL SYSTEM DESIGN FOR WATER QUALITY MANAGEMENT

5.1 Introduction

In recent years mathematical modelling and control theory have been widely applied to diverse problems outside the traditional disciplines of engineering and the physical sciences. Whether in biology, economics or the environmental sciences a distinguishing feature of many such applications is that the mathematical models representing the system under study must be realistically regarded as poorly defined, either structurally, parametrically or both. In the environmental sciences, for example, these uncertainties arise because the biological processes and complex chemical reactions that take place in the natural environment are often not well understood, at least in quantitiative terms. Further, data are limited both in quantity and quality and nonstationarity is the rule rather than the exception. Nevertheless, the ultimate goal of many efforts relating to understanding and modelling environmental systems is to develop a management scheme for mitigating some particular problem. Thus, in the environmental sciences, as in other of the new areas of application of control techniques, there is a need for methods of control system synthesis or management system design which deal explicitly with uncertainty in the process model or are sufficiently robust to assure acceptable operation in spite of uncertainty. The concept of fuzzy control (Tong, 1977), for example, is in some respects a response to this need. In this section of the proceedings an approach to the design of robust controllers is presented which relies on computer simulation rather than on an analytical framework and, as such, seems to present significant practical promise. The approach is an extension of the sensitivity analysis techniques described previously and has been applied to a river water quality control problem (Section 5.2) and the closed loop control of an activated sludge treatment plant (Section 5.3). In Section 5.4 the control of on-line blending systems is considered and applied to a nitrate control problem on the Bedford Ouse.

5.2 Control of ill-defined systems - A case study on the River Cam

The generalised sensitivity analysis described and applied in earlier sections of the proceedings depends on an ability to construct plausible model structures, to estimate broad ranges of parameter values (from limited field data or from the literature), and to define, rather loosely, the system behaviour definition, is crucial to the method and it is worth emphasizing that the defining algorithm need not be analytic: thresholds, topological conditions, logical conditions, etc. are all permissible.

The essential features of the approach are based on the assumption that:

(1) the problem under investigation can be qualitatively characterized by specific patterns of system response that define the "behaviour" of concern;

(2) one or more mathematical models of the system can be developed based on the relevant physical, chemical or biological mechanisms that are assumed to underlie the problem behaviour;

(3) these models can be parameterized by statistical distributions rather than point estimates as a means of incorporating the uncertainty in the "actual" values of the parameters.

If, in a particular case, these conditions can be met it is possible to conduct a Monte Carlo simulation by randomly selecting a parameter set from the pre-defined multivariate distribution, integrating the system equations and classifying each simulation run according to the occurrence or non-occurrence of the problem defining behaviour. A repetition of this procedure n times leads to the accumulation of m parameter vectors which lead to the behavior (B) and n - m which lead to the non-behavior ($\bar{B}$). The essential idea concerns the separation of the *a priori* parameter distributions under the behavioural mapping. That is, given the *a priori* cumulative distribution for the parameter ξ_1 as $F(\xi_1)$, the issue concerns the degree to which $F(\xi_1|B)$ differs from $F(\xi_1|\bar{B})$. Clearly, if $F(\xi_1|B) = F(\xi_1|\bar{B}) = F(\xi_1)$ then it would seem that the parameter ξ_1 was not important in determining the occurrence or non-occurrence of the behaviour. It transpires that this is a necessary but not sufficient condition for insensitivity but together with its elaborations it is the central notion of the approach.

Most techniques that are aimed at elucidation of parametric sensitivity involve some form of linearization about a point in the parameter space. The virtue of the generalized sensitivity analysis is that it is a regional rather than a local concept. In the Monte Carlo approach the issue is whether or not a particular parameter is important to the behavioural outcome of the simulations run over the entire region of the parameter space defined by the *a priori* parameter distributions. This is a critical advantage when studying poorly-defined systems of the type described in sections 4.3 and 4.4.

There is an obvious appeal in the notion of extending the sensitivity concept to the problem of controlling systems that are parametrically ill-defined. The most straight-forward extension to the control problem is to consider the design of a controller that will deliver a high probability of adequate performance under the uncertainty in knowledge of the process parameters manifested by these *a priori* distributions. Here the binary classification notion of the sensitivity approach is retained in the form of adequate or inadequate system performance. Moreover, since this performance is to be based on the simulation results it can be defined in very practical terms and as stated previously it requires only an algorithmic definition rather than an analytically tractable formulation.

This simplest approach to controller design would appear to involve the specification of one or more candidate controller structures together with a set of control parameters for each structure. Each control parameter set would then be assigned a distribution of

allowable values, the one specific set of control parameter values that maximize the probability of adequate performance, P(B). Then, the controller structure with the highest P(B) is the best of the candidates with the particular value of P(B) allowing the designer to decide if the risk can be accepted and the design implemented or if greater knowledge of the process will be needed.

This extension of the generalized sensitivity procedure to explore the problem of control of poorly defined systems is the basis of this section of the proceedings. The development and description of the method is presented below in the context of a problem involving the control of water quality in a river.

5.2.1. Description of the problem

The problem to be analyzed is based on an investigation carried out by Young and Beck (1974). A waste stream of constant strength and constant discharge flows into a lagoon. Releases form the lagoon to the river are to be scheduled such that dissolved oxygen in the river is not driven below a specified water quality standard (the behavioural criterion in our formulation) and such that the lagoon does not overflow nor fall below some minimum level.

The model for a reach of the River Cam given by Young and Beck is in terms of dissolved oxygen (DO) and biochemical oxygen demand (BOD);

$$\frac{dx_1}{dt} = -a_1 + \frac{Q + Q_E}{V_m} x_1 - a_2 x_2 + \frac{Q}{V_m} C_1 + a_1 C_s - D_B + a_4(I_k - \bar{I}) + \frac{Q_E}{V_m} C_E, \tag{5.1}$$

$$\frac{dx_2}{dt} = -a_2 + a_3 + \frac{Q + Q_E}{V_m} x_2 + \frac{Q}{V_m} L + L_A + a_5(I_k - \bar{I}) + \frac{Q_E}{V_m} L_E. \tag{5.2}$$

A simple mass balance on the lagoon yields a third equation;

$$\frac{dV_L}{dt} = Q_L - Q_E \tag{5.3}$$

where x_1 = output (downstream) DO (mg l^{-1})
x_2 = output BOD (mg l^{-1}),
L = input (upstream) BOD (mg l^{-1}),
C_1 = input DO (mg l^{-1}),
Q_E = discharge from lagoon (m^3 day^{-1}),
Q = river discharge (m^3 day^{-1}),
V_m = mean volume of reach (m^3),
V_L = lagoon volume
C_E = effluent concentration mg l^{-1}
C_s = DO saturation level mg l^{-1}

a_1 = reaeration rate constant (day^{-1}),
a_2 = BOD decay constant (day^{-1}),
a_3 = sedimentation rate constant (day^{-1}),
L_A = mean rate of addition of BOD to the reach by local runoff ($mg\ l^{-1}\ day^{-1}$),
D_B = net rate of removal of DO from the reach due to various components of respiration ($mg\ l^{-1}\ day^{-1}$),
I_k = a "sustained sunlight" term to account for the observed correlation between sunlight and high concentrations of DO and BOD.
$\bar{I}$ = the threshold level of the sustained sunlight effect
a_4 = DO rate constant for the sustained sunlight term.
a_5 = BOD rate constant for the sustained sunlight effect
Q_L = discharge to the lagoon ($m_3\ day^{-1}$),
L_E = BOD concentration of lagoon ($mg\ l^{-1}$).

The sustained sunlight term is defined by Young and Beck as

$$I_k = I_{k-1} + \frac{1}{\tau_s} h_k \frac{(\theta_k - \bar{\theta})}{\bar{\theta}} - I_{k-1} \qquad (5.4)$$

where k = time index
τ_s = time constant of the low-pass filter (days),
h_k = period of sunlight during the k^{th} day (hrs),
θ_k = river water temperature during the k^{th} day (^oC),
$\bar{\theta}$ = a mean water temperature (^oC),

The parameters associated with the model of the DO-BOD process are listed in Table 5.1 together with the values reported by Young and Beck (1974).

The controlled variable in this problem is Q_E, the lagoon discharge. The value of this variable was determined from a control law which uses state variable feedback on output DO and BOD and integral error on DO (Young and Beck,1974). The control law is:

$$Q_E = k_1(x_1 - y_r) + k_2(x_2 - \bar{x}_2) + k_3 x_3 \qquad (5.5)$$

where y_r = dissolved oxygen set point;
$\bar{x}_2$ = mean BOD level in the stream;
k_1, k_2 and k_3 are control parameters,

and x_3 is the integral error variable and defined by a third state equation:

$$\frac{dx_3}{dt} = x_1 - y_r$$

which was introduced by Young and Beck in order to control x_1 to the desired set point, y_r. In our treatment we modified this equation such that only values of x_1 less than y_r are of concern and:

$$\frac{dx_3}{dt} = x_1 - y_r \qquad x_1 < y_r$$

$$\frac{dx_3}{dt} = 0 \qquad x_1 > y_r$$

Also when x_1 rises above y_r, x_3 is reset to zero.

Table 5.1

Parameter values as given by Young and Beck (1974)

Parameter	Numerical Value
a_1	0.2
a_2	0.32
D_B	0.5
a_4	0.31
a_5	0.32
C_A	2.0
L_E	20.0
$\bar{I}$	6.0
τ_s	4.0
V_m	15.1×10^4
Q_L	2.8×10^4
$\bar{\theta}$	8.0

The method for examining control strategies is similar to that for performing a generalised sensitivity analysis. A range of possible values for the process parameters listed in Table 5.1 is chosen to reflect the system uncertainty. In this instance we simply used a rectangular probability density with a range of ± 25% of the listed values to characterise the process parameter distributions. An example of how these distributions are specified in practice is contained in Hornberger & Spear (1980). A broad range of values was used for the parameters of the control law: the *a priori* distributions for k_1, k_2, and k_3 were bounded by $[-2.0 \times 10^5, 0]$, $[0, 1. \times 10^5]$ and $[0, 1. \times 10^5]$ respectively. The distribution bounds on the set point parameters, y_r and $\bar{x}_2$, were taken to be [5.0,8.0] and [5.0, 8.10] respectively. The stream standard for DO was taken as 5.0 mg/l. As indicated above this value constitutes the behavioural definition; if during a simulation run x_1 goes below 5.0 mg/l the run is a non-behaviour and conversely. The input data for the 80 days of each simulation run were those for the River Cam as reported by Beck (1978).

As indicated above, given the foregoing model and data, it is possible to carry out a number of Monte Carlo simulations by randomly selecting a parameter set from the pre-defined distributions, integrating the systems equations over an 80 day period and classifying each simulation run according to the occurrence or non-occurrence of the "behaviour". In order to assess the benefits of control, however, it is necessary to know the behavioural probability in the absence of control. Many systems, of course, will not operate at all without

control but environmental systems will often do so. Therefore, the first simulation runs assumed the waste stream to be discharged directly to the river. Once the probability of behaviour in the open loop $P(B_o)$, is estimated the marginal benefit of control, $M_B = P(B) - P(B_o)$, can be found for various controller designs.

5.2.2. Simulation results

The initial MonteCarlo run of 250 replications was carried out with process parameter bounds of ± 25% of the values shown in Table 5.1 and without control. Twenty behaviours occurred resulting in an estimate of 0.08 for $P(B_o)$.

A second run of 250 replications was carried out with the process parameter bounds unaltered and the control parameter bounds as given above. The ranges for the gain parameters k_1, k_2 and k_3 contain the fixed set selected by Young and Beck on the basis of desirable pole locations for the linearized system. The rationale for such wide bounds on these parameters is simply to give the analysis ample opportunity to discover those portions of the control parameter space which are particularly rich in behaviours.

Of the 250 Monte Carlo runs 85 were behaviours and 165 non-behaviours The kolmogorov-Smirnov statistics $d_{m,n} = \sup|S_n(\xi_1) - S_m(\xi_1)|$, where the $S(\xi_1)$ are the sample distribution functions of the parameter ξ_1 for n behaviour and m non-behaviours indicated that 4 of the 13 process parameter distributions and two of the five control parameter distributions separated under the behavioural mapping at above the 95% level of significance ($d_{n,n} = 0.182$). These were a_2, D_B, $\bar{I}$ $\bar{\theta}$ and the control parameters k_1 and y_r. Among these process parameters $d_{n,n}$ was approximately .200 for a_2, $\bar{I}$ and $\bar{\theta}$ and .183 for D_B, the latter value being just marginally in excess of the 95% value of .182. The $d_{m,n}$ values for the five control parameters k_1, k_2, k_3, yr and x_2 were .219, .132, .094, .320, and -.167, respectively. From an inspection of the cumulative distributions it was found that the portion of the control parameter sub-space in which a higher proportion of behaviours will be found is at the low end of both the k_1 distribution and the y_r distribution with k_2, k_3, and $\bar{x}_2$ being of little apparent consequence. The correlation matrix under the behaviour contained values generally less that 0.2 with some interesting exceptions occurring for the correlations of $\bar{\theta}$ with the control parameters; 0.29 with k_1, -.32 with k_2 and -.32 with y_r.

Utilizing the results of the previous run the bounds for k_1 and y_r were altered to $(-2 \times 10^5, -1.5 \times 10^5)$ and (5, 6) respectively. All other bounds were as before. These changes resulted in raising the behaviour probability to 54% from the original 34%. In this region of the parameter space there are some changes in the list of sensitive process parameters with D_B disappearing and a_1, L_E and V_m appearing as important as judged by $d_{m,n}$ values significant at the 95% level. Of greater interest, however, are the changes in the control parameter sensitities. As expected, k_1 and y_r disappear with $d_{m,n}$ values of 0.09 and 0.17 respectively. However, k_2 and $\bar{x}_2$ now have $d_{m,n}$ values of about 0.3. This result and the low correlation between "DO" and "BOD" control parameters suggests that this region of space is good as far as the DO component of the controller is concerned, but that further improvements are possible in the BOD component ie k_2 and $\bar{x}_2$. As before, k_3 appears to be of little importance which is

not surprising in view of our behavioural definition. Hence, setting $k_3 = 0$ leads to a desirable simplification.

The distribution $S_n(k_2|B)$ indicates that behaviours are preferentially associated with values of k_2 on the low end of the range. As before, we might alter the k_2 distribution to cover the bottom 25% or so of the present range. Alternatively, we might recognize the practical fact that BOD analysis takes 5 days to accomplish and a scheme using x_2 feedback is not feasible for real time control purposes. The latter course leads to a choice of $k_2 = 0$.

These choices simplify the design problem to that of choosing specific values for k_1 and y_r from the narrowed ranges given above. However, since neither $S_n(k_1)$ nor $S_n(y_r)$ separate under the behavioural mapping, the strategy used to arrive at the present region provides little further guidance. That is, it is knowledge of the fine structure of $F(k_1|B)$ and $F(y_r|B)$ that is necessary to obtain further information. To obtain such information is costly in terms of computer time since large numbers of replications are required to obtain a good picture of the details of F from S_n. Before proceeding further on the practical questions, let us indicate what could be done with a good estimate of F if it were available.

We are seeking regions of the control parameter sub-space in which the probability of behaviour is high or, conversely, the probability of non-behaviour is near zero. Let U be the event that the m dimensional control parameter vector lies within a bounded region such that $a_i \leq k_i \leq b_i$ for i = 1,m where a_i lie within the limits of the rectangular distribution defined for k_i. Then

$$P(\bar{B}|U) = \frac{P(\bar{B})\,P(U|\bar{B})}{P(U)}$$

where $P(\bar{B})$ is the probability of non-behaviour in the entire parameter space as originally defined. Since we assume each element of the control vector $\underline{k}$ is independently distributed, $P(U|\bar{B}) = \prod_{i=1}^{m} P(U_i|\bar{B})$ where U_i is the event that $a_i \leq k_i \leq b_i$. Likewise $P(U) = \prod_{i=1}^{m} P(U_i)$. Since we desire to find the region U such that $P(\bar{B}|U) = 0$ this is equivalent to finding regions where $\prod_{i=1}^{m} P(U_i|\bar{B}) = 0$. To locate this region we may inspect the cumulative distribution of each of the k_i under $\bar{B}$ since $P(U_i|\bar{B}) = F(b_i|\bar{B}) - F(a_i|\bar{B})$. That is, we are looking for "flat" places on each of the functions $F(k_i|\bar{B})$.

If the behavioural mapping had resulted in appreciable covariance among the elements of $\underline{k}$ ($\bar{B}$) an analogous argument can be developed for dealing with a parameter set transformed by the matrix which diagonalizes the covariance matrix $E(\underline{k} - \underline{\mu})(\underline{k} - \underline{\mu})^T$ where $\underline{\mu} = E(\underline{k}|\bar{B})$

Returning to the practical issue one can either use $S_n(k_i|\bar{B})$ and $S_n(y_r|\bar{B})$ for n = 250 and assume that any flat spots are real or carry out further runs to increase n. The choice clearly depends on the cost/benefit situation for the problem at hand. Here we take the former course and select $k_1 = -1.8 \times 10^5$ and $y_r = 5.4$. These values

together with $k_2 = k_3 = 0$, resulted in an estimate of the behavioural probability of 0.84. We cannot contend that this is the best that can be done but it is a design which raises the probability of keeping the DO in the stream above 5 mg/l from 0.08 without control to 0.84 with a particularly simple control scheme.

Under the above design conditions the occurrence of the behaviour is sensitive to five process parameters. They are a_1, a_2, $\bar{I}$, V_m and $\bar{\theta}$. Of these, $\bar{\theta}$ is overwhelmingly important with a $d_{m,n}$ of 0.687 as contrasted with the 95% value of 0.232. In fact, $S_n(\bar{\theta}|B)$ is zero until $\bar{\theta}$ is near its mean value. Hence, if in the real system the value is $\bar{\theta}$ is low, behaviour will be obtained with a probability very near to unity with this controller design.

The final result, then, is that with the defined inputs and the process parameter distributions the desired behavior of the controlled process can be achieved with a probability of 0.84 with a particularly simple controller design. If this probability were deemed insufficient three courses of action are open, at least in theory; one might enlarge the lagoon thereby increasing its buffering capacity, one might engage in further research to narrow the process parameter uncertainty, or one might investigate other controller structures. In the case of process uncertainty, the results of the analysis strongly suggest that algal photosynthetic activity is critical in that the distribution of the parameter $\bar{\theta}$ separated with a $d_{m,n}$ of 0.687 in the final Monte Carlo run as discussed above. The related parameter $\bar{I}$ also was marginally important.

The possibility of increasing the lagoon volume was considered briefly. A run of 100 replications was carried out with a lagoon volume corresponding to a 25 day detention time as opposed to the original 15 days. Eighty eight behaviours were observed which did not appear to be a dramatic improvement over the 84% observed with the 15 day lagoon. Again, if trade-off decisions were to be made on the basis of these figures attention would have to be paid to the accuracy of the estimates of behavioural probability. It is clear, however, that the method yields quite useful data on the potential effectiveness of the alternative approaches to increasing the behavioural probability.

5.2.3. Discussion

The clear result of the foregoing analysis is that it was possible to determine a robust controller, that is, a controller of defined structure with fixed parameters that resulted in a relatively high probability of achieving the desired system performance in the presence of uncertainty in the process parameter values. Clearly, our success in this case rests on the fact that there was not too great an uncertainty in the process parameters, a circumstance that would nonetheless have been obvious from the analysis. In that regard it is instructive to consider the effect of process parameter uncertainty on the control parameter distributions under $\bar{B}$ and B. Figure 5.1 shows a scatter diagram of the normalized values of k vs. y_r where their distribution limits were $(-2 \times 10^5, 0)$ and (5.0, 8.0) and for which $k_2 = k_3 = k_5 = 0$ with the ± 25% variation on the process parameters. Clearly, behaviors are associated with low values of y_r with some tendency to be associated with lower values of k_1 as well. There

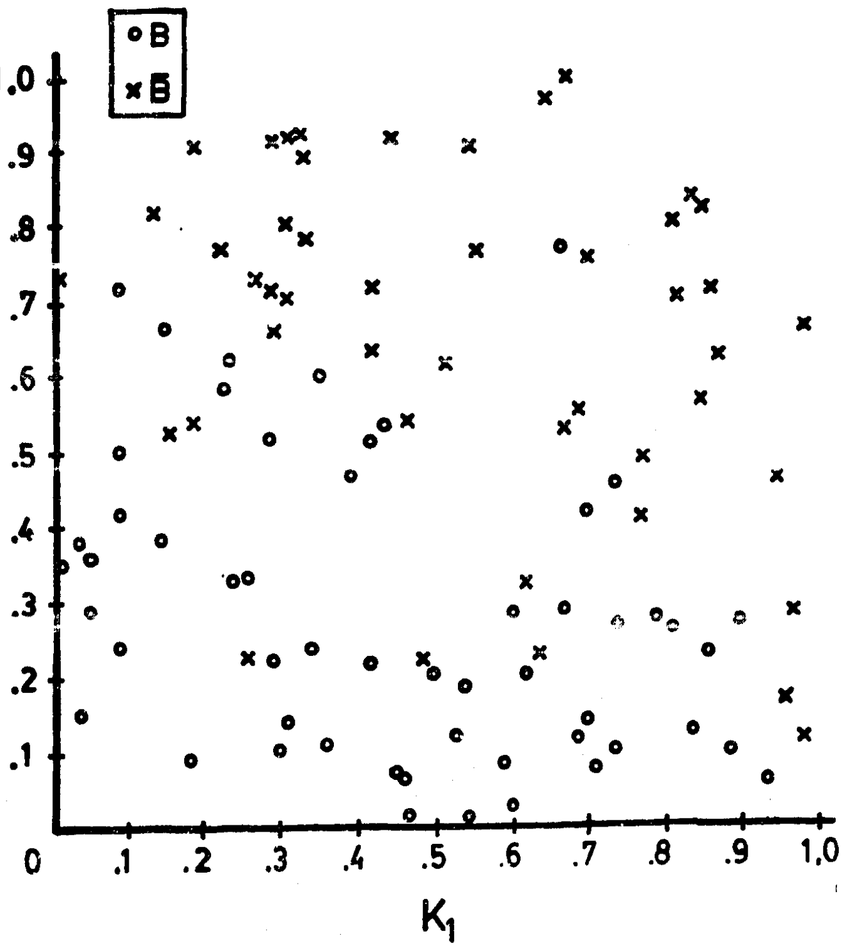


Figure 5.1 Normalized values of k_1 vs. y_r with random variation of process parameters within ± 25% of mean values

are, however, some nonbehaviours mixed in, even in the region where behaviours are most dense. In contrast, Figure 5.2 shows the same scatter diagram for identical bounds on the control parameters but with the process parameters held constant at their mean values. Here, the separation is complete and without ambiguity. In the latter case, we are no longer dealing with probabilities but merely determining controller parameters for a perfectly defined process, an interesting possibility which is discussed further below. It is clear that the opposite extreme is also possible in which the uncertainty in the process parameters is sufficiently great to overwhelm the controls and no separation will be seen in the control parameter statistics. In the latter case process identification and parameter estimation studies would be required. In such cases, however, the results of an analysis of the sort presented here will give some insight into the critical subprocesses within the system which should be the focus of special attention.

As suggested in the discussion of Figure 5.2, the Monte Carlo procedure can be used for the design of controllers for well defined processes and it is particularly attractive where non-analytic performance criteria are desirable or in cases where the process is nonlinear or otherwise analytically intractable. This aspect of the approach has been dealt with elsewhere (Auslander et al., 1981) but to mention it serves to emphasize that the simulation approach that we have proposed here is a viable, indeed an eminently practical, if somewhat inelegant alternative to the sophisticated analytic approach to control systems design that has dominated the literature for the last several decades. The method proposed herein is not a well defined and specific procedure for control system design but, rather, an experimental approach which utilizes the computer as well as techniques of statistical inference. It would appear that the general approach can be modified and elaborated to address a wide variety of practical problems.

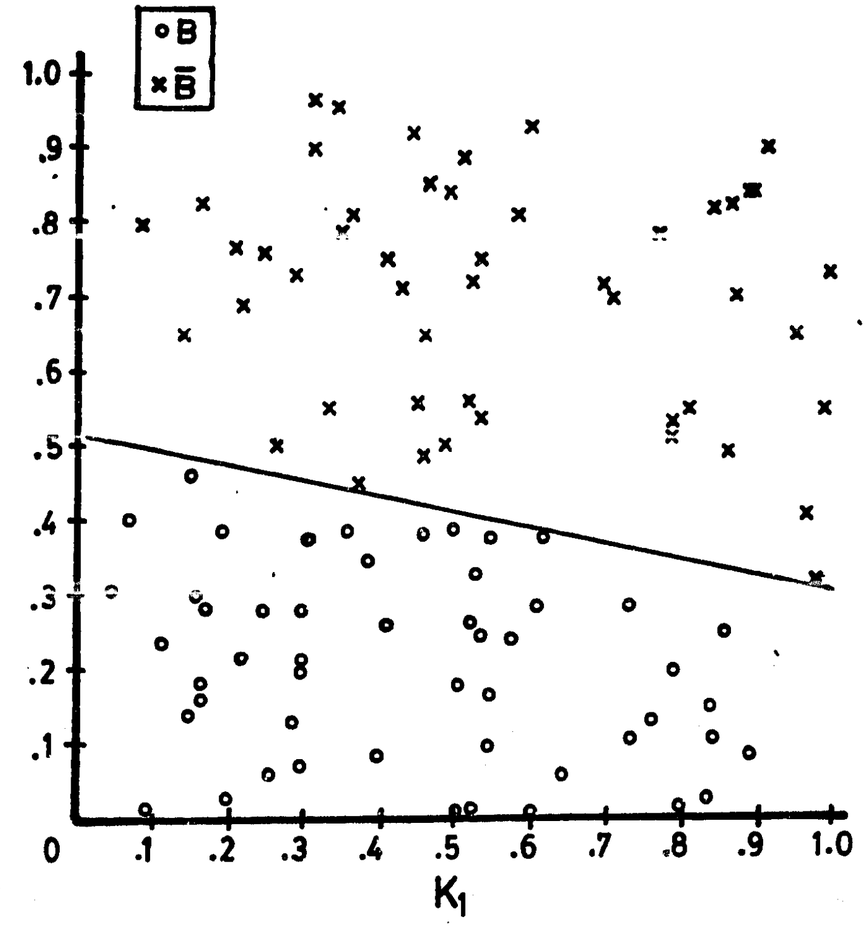


Figure 5.2 Normalized values of k_1 vs. y_r with process parameters constant at mean values.

5.3 Control of the Activated Sludge Process

As a second example of the use of the generalized sensitivity analysis to design a robust controller, the control of wastewater outflow from an activated sludge treatment plant is considered. The simulation model and associated process parameters were discussed in Section 4.4 of this report. In this section, we describe the control systems and perform the generalized sensitivity analysis using the process parameters that were judged to be "open loop consistent" (see Section 4.4).

5.3.1 Control Laws

Control actions can be undertaken by adjusting the blower volumetric flow rate (u_a), the recycle rate (r), and the sludge wastage rate (w). These options are used in the analysis to control dissolved oxygen and MLSS in the aerator. The equations describing the control actions and the associated control parameters are given in Table 5.2.

The blower flow rate is used to control dissolved oxygen levels in the aerator. This control is provided in our model through a high gain feedback loop that acts to change U_a whenever the measured value of dissolved oxygen (X_6) deviates - or tends to deviate - from the set point value (DO_{sp}). The design involves an integration in the forward loop and proportional plus derivative action in the feedback loop. A high forward loop gain results in moving the dominant pole of the linearized system to the left thereby substantially increasing the response time of the oxygen loop. The forward loop gain is the product of the controller gain and the variable process term $(O_s - X_6)$. Insofar as we wish to maintain the forward loop gain at a relatively constant value it was assumed that $(O_s - X_6)$ is measurable and an adaptive gain changer was used to switch the control gain based on the value of $(O_s - X_6)$. The resulting control law and control parameters are shown in Table 5.2.

The recycle rate is used for short term ($\sim$ hourly) control of MLSS in the aerator. A common strategy for using the recycle rate to control MLSS is "ratio control" in which the recycle rate is set to a proportion of the influent flow rate. We adopted a variant on this idea by including a derivative term in the control law (see Table 4.2) to allow for the potential of more rapid response to influent changes.

The sludge wastage control law is determined to provide for longer term ($\sim$ daily) control of MLSS contingent upon the sludge blanket level being lower than 90% of its maximum value. A simple proportional control using a daily average set point value of MLSS in the aerator was used. The wastage rate affects the amount of sludge in the settler which in turn determines the concentration of MLSS in the return flow. Thus, changes in wastage rate will affect the long-term equilibrium value of MLSS in the aerator. Our control is thus similar in aim to that discussed by Tanuma (1980). Table 5.2 contains the control law and related parameters.

The ranges of values for the control parameters chosen for the Monte-Carlo simulations are given in Table 5.3. These were set to give a

TABLE 5.2 Control laws for the closed-loop analysis of the activated sludge process.

Dissolved Oxygen

$$\dot{u}_A = \text{GAIN}\ (DO_{sp} - X_6 - \dot{X}_6)$$

where GAIN is defined by:

$(O_x - X_6) \geq \varepsilon$, GAIN = K_1

$(O_s - X_6) < \varepsilon$, GAIN = K_2.

Control parameters: DO_{sp}, ε, K_1, K_2.

Restrictions: $0 \leq u_a \leq 20000\ m^3\ hr^{-1}$

Recycle Rate

$$r = \lambda_1 Q_I + \lambda_2 \dot{Q}_I$$

where $\dot{Q}_I \equiv dQ_I/dt$.

Control parameters: λ_1, λ_2

Restrictions: $0.1 \leq r \leq 2.0$

Wastage Rate

$$w = N(\overline{MLSS} - MLSS_{sp}) + w_o$$

where $\overline{MLSS}$ is the daily average value of suspended solids in the aerator.

$w_o = w_{MAX}/2$

$N = w_{MAX}/2000.$

Control parameters: w_{MAX}, $MLSS_{sp}$.

Restrictions: If $x_8 \geq 0.9\ V_c$, $w = w_o$

TABLE 5.3. Values of control parameters for the Monte-Carlo simulations.

Parameter	Range
DO_{sp}	1 - 3.5
ε	1 - 20
K_1	1000 - 9000
K_2	5000 - 20000
λ_1	0 - .001
λ_2	0 - .007
w_{MAX}	... - .14
$MLSS_{sp}$	1000 - 5000

broad range of values of the control variables u_s, r and w for conditions specified to be generally consistent with the Norwich plant.

5.3.2. Simulation results

The 354 sets of process parameters that were determined to be "open loop consistent", as discussed in Section 4.4, were used in the simulations with the control parameters chosen from a uniform distribution with the limits given in Table 4.3. The behaviour criteria for the closed loop were changed to reflect a desired reduction in the DO concentration variation in the aerator and in the concept of NH_3 and BOD, the two most commonly-used indices of effluent quality. Adequate behaviour in the controlled case was considered to have been achieved only if the concentration of NH_3 in the effluent did not exceed 15mg l^{-1}. DO was required to remain between 0.5 and 4 mg l^{-1}. Criteria for adequate behaviour as judged by values of other state variables remained unchanged from the open loop analysis. Of the 354 simulations that were open loop consistont, 112 satisfied the more stringent closed loop behaviour criteria. The probability of obtaining closed loop behaviour in the open loop (without control) is thus 32%.

The 354 simulations with the wide bounds on control parameters listed in Table 5.3 resulted in only 65 behaviours. Essentially no behaviours occurred for values of the control parameter λ_1 <.0005. Consequently a second set of simulations was run with the bounds on λ_1 changed to 0.0005 to 0.001 and the bounds on λ_2 changed to 0 to 0.005. For these modified bounds 128 behaviours were obtained. Table 5.4 lists the values of the Kolmogrov-Smirnov statistic for Case I, the "modified wide bounds case". This is again a two-sample statistic for comparing parameter differences between the behaviour class and the non-behaviour class. The statistic for seven of the process parameters is significant at above the 99% level. These are associated with nitrifier growth/NH utilization (Y_1, $\hat{\mu}$, K_{m6}), with MLSS growth/BOD utilization (Y_2, $\mu_m{}^3$, δ) or with oxygen transfer (k_a). Among the control parameters, DO_{sp} and λ_1 were of considerable importance in distinguishing behaviours according to the statistical analysis. To further refine the range of desirable control parameters another set of simulations was run with the range of DO_{sp} set to 2.5 to 3.5 and the range for λ_1 set to 0.00075 to 0.00125. Statistics for this run are listed under Case II in Table 5.4. The number of behaviours for Case II was 1[illegible] probability of behaviour with control, $P(B_{CL})$, of 54%.

These results suggest that behavioural out[illegible] sensitive to a number of process parameters and that satis[illegible]trol may be difficult to achieve in the face of significant un[illegible]ainty in their values. Examination of the cumulative distributions of parameters under the behavioural classification suggested that some improvement

TABLE 5.4 Kolmogorov-Smirnov statistic

Parameter	CASE I Modified wide bounds 99% value = .180	CASE II Second modification of control parameter bounds 99% value = .174	CASE III Control design 99% value = .176
Process Parameters			
Y_1	.212	.237	.226
$\hat{\mu}$	.299	.365	.460
K_{m_1}	.104	.110	.111
k_4	.117	.174	.152
μ_m	.230	.175	.191
b	.080	.062	.102
K_{m_4}	.091	.147	.140
k_a	.183	.183	.161
α	.117	.090	.076
δ	.183	.150	.198
p	.149	.089	.103
θ	.068	.057	.054
K_{m_6}	.187	.211	.200
Y_2	.185	.136	.081
X_m	.150	.104	.050
Control Parameters			
DO_{sp}	.239	.086	--
ε	.087	.096	--
K_1	.050	.067	--
K_2	.142	.066	--
λ_1	.215	.114	--
λ_2	.160	.110	--
w_{MAX}	.109	.093	--
$MLSS_{sp}$	.128	.077	--

in $P(B_{CL})$, however slight, might be realized by choosing a set of "design" control parameters. The design set of parameters is listed in Table 5.5. Larger values of DO_{sp} were clearly favourable for behaviours. However, very high set point values are unrealistic and 3 mg l^{-1} was chosen as a compromise. The adaptive gain change in the DO controller was not substantially important in determining behaviour so the DO controller was simplified to a single gain of 10,000. Values of λ_1 in the mid to upper range were indicated and a value of 0.0011 was selected. A relatively flat portion of the cumulative distribution of λ_2 in the non-behaviour category appeared to occur near the selected value of 0.0035; as indicated in a previous section of this report, such flat portions are likely places for selecting "good" design parameters. Finally, somewhat large values were indicated for both w_{MAX} and $MLSS_{sp}$ and the design values of 0.10 and 4000 mg l^{-1} reflect this.

Table 5.5: Design values of control parameters

Parameter	Value
DO_{sp}	3
K_1, K_2	10000
λ_1	0.0011
λ_2	0.0035
w_{MAX}	0.10
$MLSS_{sp}$	4000

With the design values of the parameters 209 behaviours we obtained for a $P(B_{CL}) = 0.59$. Ninety-one of the behaviours were due to violations of the NH_3 bounds, 49 were due to violations of BOD bounds, and 5 to violations of the DO bounds. The Kolmogorov-Smirnov statistic for the process parameters in the design case are listed in Table 5.4. Five of the original seven important process parameters, as judged by the 99% confidence level, remained so in the design run. The two process parameters that dropped in importance ranking are associated with the dissolved oxygen dynamics (k_a) and with BOD dynamics (Y_2). Evidently control of DO with the set point at 3 mg l^{-1} reduces k_a in importance in terms of simulation of adequate performance and recycle and/or wastage control reduces the importance of Y_2. However, the overwhelming importance of uncertainty in the process parameters is clearly indicated by the values of $d_{m,n}$ for the other five parameters, particularly that for $\hat{\mu}$.

We define the probability of obtaining closed loop behaviour in the open loop, $P(B_O)$, with reference to the original simulations (see Section 4.4). That is, fixed values of U_a, r and w are chosen at random along with values of the process parameters; no attempt was made to determine "optimal" control settings for a given set of values of process parameters. Using this definition of $P(B_O)$, the marginal benefit of control is calculated as $P(B_{CL}) - P(B_O)$ where $P(B_{CL})$ is the probability of satisfying the closed loop criteria using control. The marginal benefit of control in this instance is only 27%. However, 91 of the 161 non-behaviours in the design run were due to NH_3 violation and none of the control actions was directed specifically toward the nitrogen cycle. Any benefits gained in control of NH_3 had to stem indirectly from the controls on MLSS. Without the closed loop criterion on NH_3, the behavioural probability of successful control would approach 85%, a much more respectable value.

The failure to achieve a high marginal benefit of control in the face of important process parameter uncertainty suggests two alternatives for improvement, neither of which offers much for pre-existing plant. The first is a research approach: studies could be undertaken to reduce the uncertainty in several key parameters. This is the approach first outlined by Spear and Hornberger (1980). The second approach is to utilize the generalized sensitivity analysis as an integral part of the plant design itself. In the present instance, for example, one might conclude that a larger aeration tank is called for if the uncertainty in parameters describing bacterial growth kinetics cannot be reduced. An experiment to demonstrate the interplay between our analysis and plant design was run to illustrate the point. The set of simulations for the design case was rerun with the value of V_A increased by 25%. The probability of behaviour with this design modification rose to 76%. NH_3 violations still comprised roughly three quarters of the non-behaviours with the remainder being mainly BOD violations. The importance of the nitrification dynamics was reflected in the statistics by the overwhelming importance of the maximum growth rate of nitrifiers which had a $d_{m,n}$ value of 0.814.

5.3.3 Discussion

The application of the generalized sensitivity analysis to a simulation model aimed at the control of an activated sludge plant reinforces the power of the methodology for examining poorly defined systems. The results for this problem indicate that a high probability of acceptable control cannot be achieved in the face of unresolvable uncertainties in the process parameters, at least with the type of controllers used. Given the crude approximations already involved in the lumped parameter model used here, it does not seem likely that a satisfactory control law is attainable for the particular structure evaluated.

The ultimate failure of the control implemented in the model was an inability to inject a high enough level of biomass into the aerator at critical times to assure proper operation in terms of BOD reduction and nitrification. With the current model structure this is

accomplished by adjusting the recycle rate. While this approach does result in short term increases in active biomass in the aerator, the longer term effect is a decrease because of the direct interconnection with the clarifier where concentrations are reduced with increasing recycle rate. Increasing the size of the aerator partially alleviates this problem by raising the detention time, thereby allowing more time for MLSS and nitrifier growth in the aerator itself. The efficacy of this design modification was demonstrated above where it was shown that the marginal benefit of control was increased from 27% to 44% by a 25% increase in aerator volume. An alternate solution to the problem would be to use a sludge storage tank which could be used to increase the active biomass in the aerator as needed. This type of approach to controlling the activated sludge process is discussed by Busby and Andrews (1975), Stenstrom and Andrews (1979) and Tanuma (1980). In the case of the activated sludge analysis, the uncertainties in process parameters are severely limiting. While the generalized sensitivity procedure might be useful in assigning priorities to some areas of needed basic research (eg see previous sections of these proceedings), much more would be required in terms of increasing the sophistication of the actual model used here before any specific application to a given treatment plant would be warranted. The analysis procedure can, of course, be applied to any model structure and interconnection with plant design can be implemented regardless of the chosen structure.

5.4 On-line blending for nitrate control in water supply

5.4.1 Nitrates and their control in water supply

The concern over the presence of high nitrate concentrations in water supply stems from its connection with infant methaemoglobinaemia. Whilst the reported cases attributed to nitrate in drinking waters are fairly few in number, the WHO has recommended stringent levels for acceptable nitrate concentrations in water supply. It would be true to state that these levels do not purport to represent any real or well-defined dose-response relationship. They are probably very conservative and reflect the attitude of modern society, which seems to be extremely averse to risks under the control of public authorities. Additionally, recent evidence has also suggested some connection of the intestines resulting from nitrosamines formed through bacterial action on ingested nitrate.

The strict standards relating to nitrate and the general increases of nitrate in lowland river sources now require some remedial actions to be initiated. The increase in nitrate concentrations stems from a variety of sources. Agricultural practice has led to increased groundwater and runoff concentrations. High ammonia loads to treatment works either lead to increased levels of nitrate in effluent discharges or, alternatively, increased in-stream nitrification, which has further effects on the 'health' of the water body. Means of dealing with high nitrate concentrations need to be developed and implemented. Furthermore, it should be noted that nitrate problems are not present all the time; they can be very transient in character and methods of dealing with these difficulties must be flexible and cost effective. It is also as well to ensure that any proposed solution is not too parochial in outlook. In some cases the problems at a downstream abstraction may be alleviated by appropriate action at points of discharge upstream, for example by operating wastewater treatment plants in a nitrifying-denitrifying mode. In cases where river nitrates come predominantly from runoff or groundwater/baseflow sources, such a solution may not offer any real improvement and some form of nitrate control must be exercised at the water supply abstraction. If the river forms the sole supply then biological nitrification using methanol, or ion exchange may be necessary. There are, though, some cases where a variety of sources is available and sensible nitrate concentrations may be achieved through a blending operation (Green, 1978).

The adopted form of nitrate control/removal must depend on the relative costs and efficiencies, and these are inextricably bound up with the temporal pattern of the nitrate concentrations. The magnitude, frequency and duration of nitrate transients may be decisive factors in discriminating between different methods of nitrate removal. The provision of pre-treatment bankside storage has been suggested as a useful buffer against quality and supply/demand transients. Although there may be some problems associated with algal growth and thermal stratification, this physically-based solution may be more acceptable than nitrate removal using methanol. In the German Federal Republic there is a trend towards providing 7 days bankside storage, a similar move is unlikely to be adopted so widely in the UK. The benefits of a small buffering lagoon will be examined below using a hypothetical system based loosely on the Clapham abstraction on the Bedford-Ouse. (See Appendix 1 for details of this system). The configuration allows river water, groundwater and reservoir water to be blended in a small lagoon prior to treatment.

The smallness of the existing lagoon (typically giving a retention time of about 18 hours) appears to offer little buffering capacity to sustained transients in abstracted river waters. Measures to counteract such quality variations consist of

(a) Increasing the lagoon volume.

- this involves high capital cost and may worsen any algal or thermal stratification problems.

(b) Augmenting the limited buffering capacity with on-line nitrate removal (eg using ion exchange or methanol)

- once again capital costs are incurred. The nitrate removal facility must also be capable of being switched on or off in response to river quality.

(c) Controlling the blending operation and bringing in greater proportions of the high quality sources in response to adverse river quality. There are several degrees of sophistication that such blending may assume. The blending could be controlled in real-time on an almost continuous basis, or be reset periodically; the greater the frequency of changes to blending proportions then the more rapid the response to nitrate transients. The costs incurred will consist of limited capital expenditure on monitoring/control equipment and variable speed pumps.

The choice of any o combination of solutions will depend on the capital and non-cap sts and the improvement in quality. Inevitably this latter aspect co ns an element of subjectivity.

5.4.2 System configuration

The system under study is shown in Figure 5.3. Waters from three sources (river, aquifer and reservoir) are mixed in a small lagoon prior to water treatment. The nitrate concentrations of each source are u, n_a, n_r respectively, abstracted at rates q, q_a, q_r to satisfy a demand Q_D. The maximum river, aquifer and reservoir abstractions are taken as Q_D, Q_a, Q_r. The lagoon contents of volume V are assumed perfectly mixed and at a concentration x_1.

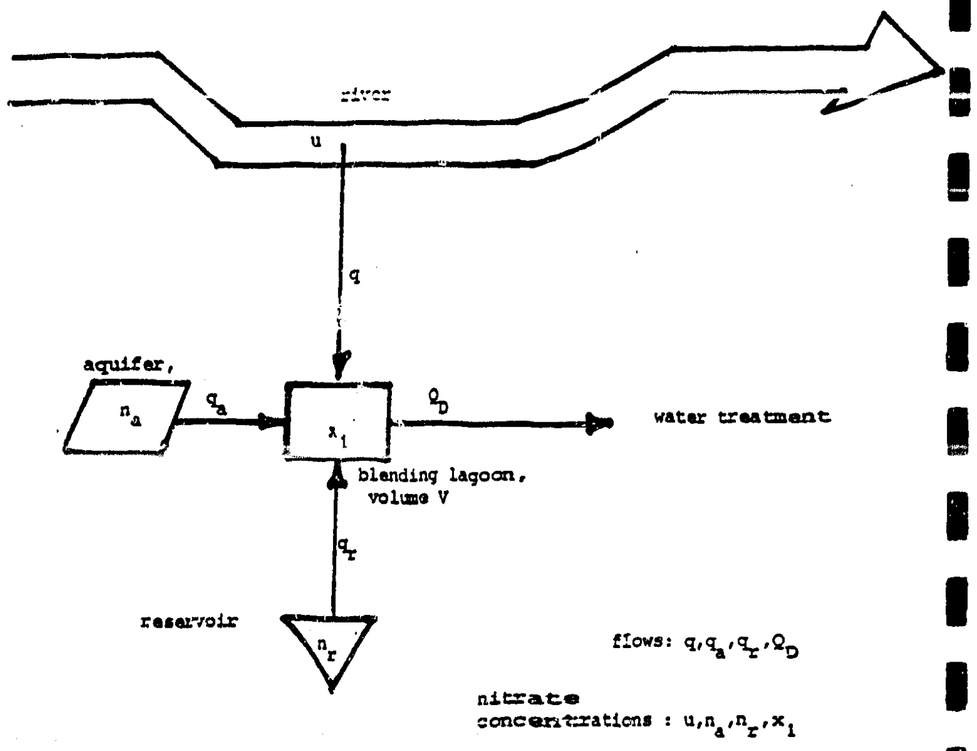


Figure 5.3 Schematic of system under study

The system dynamics are summarised by the mass balance equations:

$$\dot{V} = q + q_a + q_r - Q_D \tag{5.6}$$

$$\frac{d}{dt}\{x.V\} = qu + q_a n_a + q_r n_r - Q_D x_1 \tag{5.7}$$

or, equivalently,

$$\dot{x}_1 = \frac{1}{V}\{qu + q_a n_a + q_r n_r - (q + q_a + q_r)x_1\} \tag{5.8}$$

Equations (5.6) and (5.8) constitute a pair of simultaneous, first order, non-linear differential equations. The abstraction rates, represent decision variables and the nitrate concentrations u, n_a, n_r constitute time-varying inputs, though n_a and n_r will vary more slowly than u. Note that if the lagoon is maintained at constant volume then equation (5.8) becomes linear.

The blending operation should achieve satisfactory quality but at minimum cost. The seasonal nature of the input nitrates will therefore lead to a corresponding pattern in abstraction rates. In addition, however, there may be sudden changes in the river nitrate level and it is necessary to examine the response of the system to such impulses and where necessary to suggest some form of control, either by a change in the blending operation or by auxiliary treatment.

5.4.3 Steady-state solutions

Whilst a steady-state condition will never exist it is nonetheless useful to examine some typical and critical conditions that may crystallise the extent of any nitrate problem. Setting equations (5.6) and (5.8) to zero gives

$$q + q_a + q_r = Q_D \tag{5.9}$$

$$qu + q_a n_a + q_r n_r = Q_D x_1 \tag{5.10}$$

and the disappearance of V allows the lagoon contents to be set quite arbitrarily.

The nitrate concentration x_1 going to supply can then be determined from the input and decision variables. Because of differences in the abstraction costs of each source it is appropriate to set an upper limit x_1^* and then seek values of the decision variables which minimise abstraction costs. The costs of each source can reflect factors other than pumping costs alone. For example, excessive aquifer drawdown can be discouraged (rather than prohibited) by including penalty costs for aquifer withdrawals. If it is assumed that river water is 'free' and the costs of the other two sources are proportional to their abstraction rates, then the minimum cost solution for a particular steady-state will be given by the following Linear Program (LP):

Minimise $z = kc_r q_a + c_r q_r$

subject to:

$$q + q_a + q_r = Q_D$$

$$qu + q_a n_a + q_r n_r \leq Q_D x_1^* \qquad \text{(LP1)}$$

$$0 \leq q \leq Q_D$$

$$0 \leq q_a \leq Q_a$$

$$0 \leq q_r \leq Q_r$$

In LP1 the decision variables are q, q_a, q_r; c_r is the cost of reservoir abstraction and k is a constant reflecting the relative costs of aquifer and reservoir abstractions.

Whenever the river nitrate concentration u is at or below x_1^*, naturally all supply can be taken from the river. If $u > x_1^*$ then some blending is necessary to keep $x_1 \leq x_1^*$. The structure of the LP will always lead to $x_1 = x_1^*$ at optimality. With this in mind, the LP is written in "Big M" or Charne's Multiplex form as follows;

minimise $z = kc_r q_a + c_r q_r + Ma + Mb$

subject to

$$a + q + q_a + q_r = Q_D$$

$$b + qu + q_a n_a + q_r n_r = Q_D x_1^*$$

$$c + q = Q_D$$

$$d + q_a = Q_a$$

$$e + q_r = Q_r$$

where a, b are artificial variables;

c, d, e are slack variables;

M is a large positive number.

The simplex solution given herein is for the case $u > x_1^* > n_r > n_a$. If $u >> x_1^*$ then an infeasible solution results requiring relaxation of the quality constraint by increasing x_1^*. If the dynamics of the system can be ignored then the simplex solutions can be used in a variety of ways for nitrate control. At the simplest level average values for u, n_a, n_r could be used, say for each month. However, this affords

little protection of the supply from quality transients or deviations from assumed monthly averages; if the lagoon is small then pulses of nitrate will receive only little attenuation. At a more sophisticated level, the "steady-state" u, n_a, n_r values could be daily or moving average values. Although this would give a better response to sustained changes in inputs, transients of only a few hours duration can once again pass unnoticed and some study of the open loop dynamics is called for. If transient conditions are found to be a cause for concern then further control measures may be required. A combination of the "steady-state" LP solution with some state variable feedback control is one possibility; incorporation of river nitrate forecasts in a feedforward control scheme is also possible. The efficacy of such forms of nitrate control has then to be compared with methods which rely on auxiliary treatment.

5.4.4 System dynamics and response

Recalling equation (5.8) and keeping V constant (so that $Q_D = q+q_a+q_r$) then

$$\dot{x}_1 = \frac{1}{V}\{qu + q_a n_a + q_r n_r - Q_D x_1\} \qquad (5.11)$$

It will be useful to accompany equation (5.11) by a second differential equation which determines the integral error, x_2, between the actual level of x_1 and its desired value x_1^*.

$$\dot{x}_2 = x_1 - x_1^* \qquad (5.12)$$

Open Loop Dynamics

Consider first the system's open loop dynamics about some nominal 'steady-state' solution in response to a change δu in the river nitrate level.

The perturbation equations corresponding to equations (5.11) and (5.12) are then of the form

$$\delta\dot{\underset{\sim}{x}} = \underset{\sim}{A}\,\delta\underset{\sim}{x} + \underset{\sim}{B}\,\delta u \qquad (5.13)$$

where $\delta\underset{\sim}{x} = \{\delta x_1, \delta x_2\}^T$

$$\underset{\sim}{A} = \begin{vmatrix} \frac{Q_D}{V}, & 0 \\ 1, & 0 \end{vmatrix}$$

$$\underset{\sim}{B} = \{\frac{q}{V}, 0\}^T$$

The open loop dynamics have the characteristic equation

$$\det[\underset{\sim}{A} - s\underset{\sim}{I}] = 0 = s(s + \frac{Q_D}{V}) \qquad (5.14)$$

This gives an undesirable open loop response to a step change δu in river nitrate of:

$$\delta x_1 = \frac{\delta u\, q}{Q_D}\{1 - \exp(-\frac{Q_D}{V} t)\} \qquad (5.15)$$

Feedback Control and Associated Closed Loop Dynamics

In order to counteract the undesirable effects of transients in river nitrate levels, some form of state variable feedback control of the blending operation can be considered. Once again a nominal 'steady-state' is assumed and changes δg to the vector of steady-state abstraction rates are related to the state perturbation vector δx by

$$\delta g = \underset{\sim}{K}\,\delta x \qquad (5.16)$$

where $\underset{\sim}{K}$ is a matrix of control gains.

Reference to equation (5.8) shows that the dynamics will remain linear if the abstraction perturbations are such that

$$\delta q + \delta q_a + \delta q_r = 0$$

Accordingly, δq_a and δq_r are defined as

$$\delta q_a \triangleq g_1\,\delta q$$

$$\delta q_r = g_2\,\delta q \qquad (5.17)$$

where $g_1 + g_2 = -1$.

Thus

$$\delta g = \begin{vmatrix} \delta q \\ \delta q_a \\ \delta q_r \end{vmatrix} = \begin{vmatrix} 1 \\ g_1 \\ g_2 \end{vmatrix} [k_1, k_2] \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} \triangleq \underset{\sim}{G}\,\underset{\sim}{K}\,\delta x \qquad (5.18)$$

Reference to equations (5.11) and (5.12) shows that the closed loop perturbation dynamics will then be given by

$$\delta\dot{x} = [\underset{\sim}{A} + \underset{\sim}{F}\,\underset{\sim}{G}\,\underset{\sim}{K}]\delta x + \underset{\sim}{E}\,\delta u \qquad (5.19)$$

where $\underset{\sim}{A} = \begin{vmatrix} -\frac{Q_D}{V}, & 0 \\ 1, & 0 \end{vmatrix}$

$$\underset{\sim}{F} = \begin{vmatrix} \frac{u+\delta u}{V}, & \frac{n_a}{V}, & \frac{n_r}{V} \\ 0, & 0, & 0 \end{vmatrix}$$

$$\underset{\sim}{G} = \begin{vmatrix} 1 \\ g_1 \\ g_2 \end{vmatrix}, \qquad \underset{\sim}{K} = [k_1, k_2]$$

$$\delta u = \begin{vmatrix} \delta u \\ \delta x_1^* \end{vmatrix}$$

$$\underset{\sim}{E} = \begin{vmatrix} \frac{q}{V}, & 0 \\ 0, & -1 \end{vmatrix}$$

The input perturbation vector δu now includes a change δx_1^* to the set point. Including δx_1^* allows consideration of feedforward control to be analysed; forecasts of high levels of river nitrate may prompt a lowering of lagoon nitrate concentration in readiness.

Although equation (5.19) remains linear, its coefficients are time-varying (notably the term $(u+\delta u)/V$ in $\underset{\sim}{F}$). This means that if $\underset{\sim}{K}$ is constant and selected a priori, there is no guarantee of stability: equation (5.19) is exact and the pertubation u is not required to be small. Stability is determined by the roots of the characteristic equation

$$\det[\underset{\sim}{A} + \underset{\sim}{F}\,\underset{\sim}{G}\,\underset{\sim}{K} - s\underset{\sim}{I}] = 0$$

Nevertheless, some useful aspects of the dynamic response can be ascertained by neglecting δu in $\underset{\sim}{F}$ and using a time invariant approximation to equation (5.18).

$$\delta\dot{x} = [\underset{\sim}{A} + \underset{\sim}{D}\,\underset{\sim}{K}]\delta x + \underset{\sim}{B}\,\delta u \qquad (5.20)$$

$$\text{where } \underset{\sim}{D} = \begin{vmatrix} \dfrac{u + g_1 n_a + g_2 n_r}{V} \\ 0 \end{vmatrix} \triangleq \begin{vmatrix} f \\ 0 \end{vmatrix}$$

$$f = \frac{u + g_1 n_a + g_2 n_r}{V}$$

and other matrices are as before.

The control gain matrix $\underset{\sim}{K}$ is then chosen to give desirable closed loop dynamics and improve the response to input perturbations. The characteristic equation, $\det[\underset{\sim}{A} + \underset{\sim}{D}\,\underset{\sim}{K} - s\underset{\sim}{I}] = 0$, yields

$$s^2 + s\{\frac{Q_D}{V} - fk_1\} - fk_2 = 0 \qquad (5.21)$$

which may be written in the form

$$s^2 + 2\omega\zeta s + \omega^2 = 0 \qquad (5.22)$$

where the natural frequency ω and the damping factor ζ together characterise the damped sinusoids that comprise the closed loop dynamics. The control gains are then related to specified values of ω and ζ by equating coefficients:

$$k_2 = (\frac{Q_D}{V} - 2\omega\zeta)/f$$

$$k_1 = -\frac{\omega^2}{f} \qquad (5.23)$$

The response to an input perturbation vector δu can be expressed conveniently in terms of Laplace Transforms:

$$\Delta\underset{\sim}{X}(s) = \underset{\sim}{M}(s)\,\underset{\sim}{B}\,\Delta\underset{\sim}{U}(s) \qquad (5.24)$$

where the transform of the impulse response matrix is given by

$$M(s) = \frac{1}{(s+\omega\zeta)^2 + \omega^2(1-\zeta^2)} \begin{vmatrix} s & , & fk_2 \\ 1 & , & s+(\frac{Q_D}{V} - fk_1) \end{vmatrix} \qquad (5.25)$$

For input disturbances of the type {δu = constant, $t > 0$} with transform {constant/s}, the state perturbation transforms will be of the form

$$\Delta X(s) = \frac{As + Bs + C}{s\{(s+b)^2+a^2\}} \triangleq \frac{Ds + E}{\{(s+b)^2+a^2\}} + \frac{F}{s} \qquad (5.26)$$

with a solution in time space as

$$\delta x(t) = De^{-bt}\cos(at) + \frac{(E-Db)}{a} e^{-bt}\sin(at) + F \qquad (5.27)$$

where
$$D = A - \frac{C}{\omega_n^2}$$

$$E = B - \frac{2\zeta}{\omega_n} C$$

$$F = \frac{C}{\omega^2}$$

Thus for $\delta\underset{\sim}{u} = \{\delta u, 0\}^T$, the perturbations δx_1, δx_2 are:

$$\delta x_1 = \frac{q}{V}\,\delta u\,[\frac{1}{\omega\sqrt{1-\zeta^2}} e^{-bt}\sin(at)] \qquad (5.28)$$

$$\delta x_2 = \frac{q}{V}\,\frac{\delta u}{\omega^2}\,[(1 - e^{bt}\cos(at)) - \frac{\zeta}{\sqrt{1-\zeta^2}} e^{-bt}\sin(at)] \qquad (5.29)$$

and for $\delta\underset{\sim}{u} = \{0, \delta x_1^*\}^T$,

$$\delta x_1 = -\delta x_1^*(1 - e^{-bt}\cos at) \qquad (5.30)$$

$$\delta x_2(t) = \frac{\delta x_1^*}{\omega}\,[-2\zeta(1-e^{-bt}\cos(at)) + \frac{(2\zeta^2-1)}{\sqrt{1-\zeta^2}} e^{-bt}\sin(at)] \qquad (5.31)$$

The equations above give the essence of the closed loop dynamics, though in reality the control perturbations induced by $\underset{\sim}{K}\,\delta x$ are subject to restrictions imposed by bounds on the abstraction rates.

the necessary adjustments to the control scheme to account for boundedness of the control variable.

5.4.5. Simulation of feedback control

Figures 5.4 and 5.5 show open loop and closed loop responses for two different river nitrate [illegible] In each case the lagoon is at an initial nitrate co[illegible] mg/l. The values of u, g_1, g_2 used to evaluate f (the [illegible] control gains k_1, k_2) in equation (5.20) are 12 mg/l, - 0.[illegible] 5 re[illegible]tively.

In each figure, the nitrate disturbances last for 2 days, rising from initial values of 12 mg/l and 10 mg/l. The former situation represents a case where blending is extant prior to the disturbance whereas in the latter case the whole supply is taken from the river.

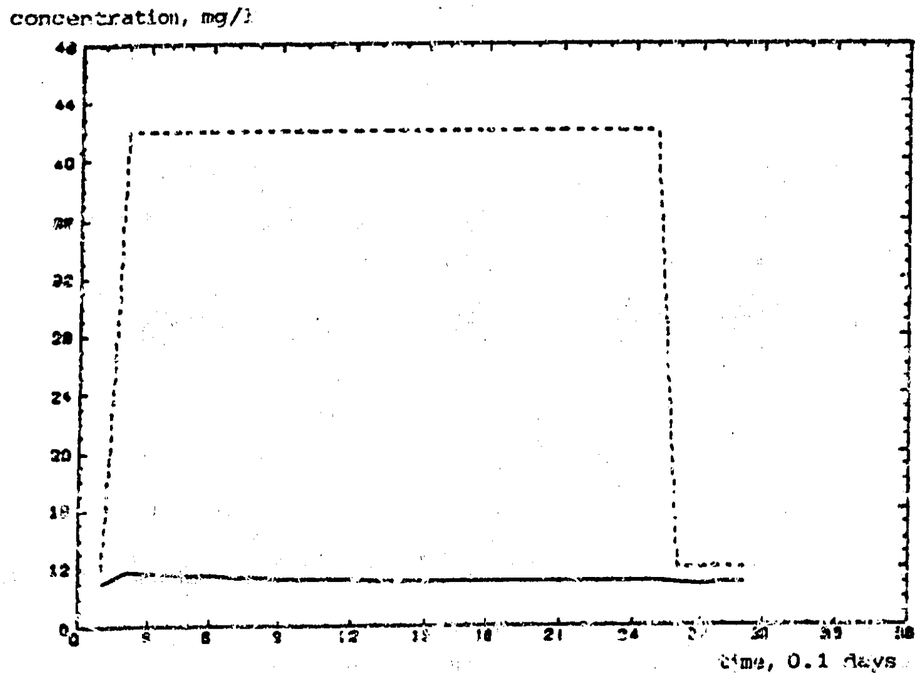
As might be expected, short term transient levels of river nitrate are almost fully attenuated by the implementation of feedback control. For "LP" control to be as effective, it would need to be initiated at very short time intervals to prevent transients from being transmitted to supply. The controller has nominal dynamics of $\omega = 6$ rads/day and $\zeta = 0.8$. In Figure 5.4 the steady state river nitrate concentration of 12 mg/l is increased instantaneously by 30 mg/l. Without feedback control, the lagoon output nitrate concentration reaches about 30 mg/l within a day. With feedback control, it never exceeds 12 mg/l. Similar effects are seen in Figure 5.5 where the initial river nitrate level is sufficiently low to allow total abstraction from the river at the outset. Without feedback control, an increment of 20 mg/l in the river nitrate level is readily transmitted to supply, reaching 70% of the increment within a day. The maximum exceedance of the standard (11 mg/l) is less than 1 mg/l with feedback control is employed.

Equations (5.30), (5.31) give the closed loop perturbations in response to changes in the set point. This might be used to draw down the lagoon nitrate level in preparation for a high shock load. The response of the feedback control scheme to high input disturbances appears good enough to render this option redundant. It would seem that feedforward action might be best suited to a control scheme implemented intermittently and not through continuous state variable feedback; LP control response to forecast nitrate loads would seem appropriate. It must be noted though that no safeguard against undetected or unpredicted nitrate transients would be offered by this approach.

5.4.6. Simulation of feedforward control

An alternative to continuous state variable feedback control is to use the LP at frequent time intervals, say daily. This has the obvious advantage of eliminating the need for sophisticated monitoring and control apparatus. The effect of undetected transients can be gauged from earlier figures.

If, though, the nitrate transients are predictable from some forecasting scheme and/or upstream monitoring station, the changes in blending dictated by daily LP control will give quite acceptable nitrate levels in supply.



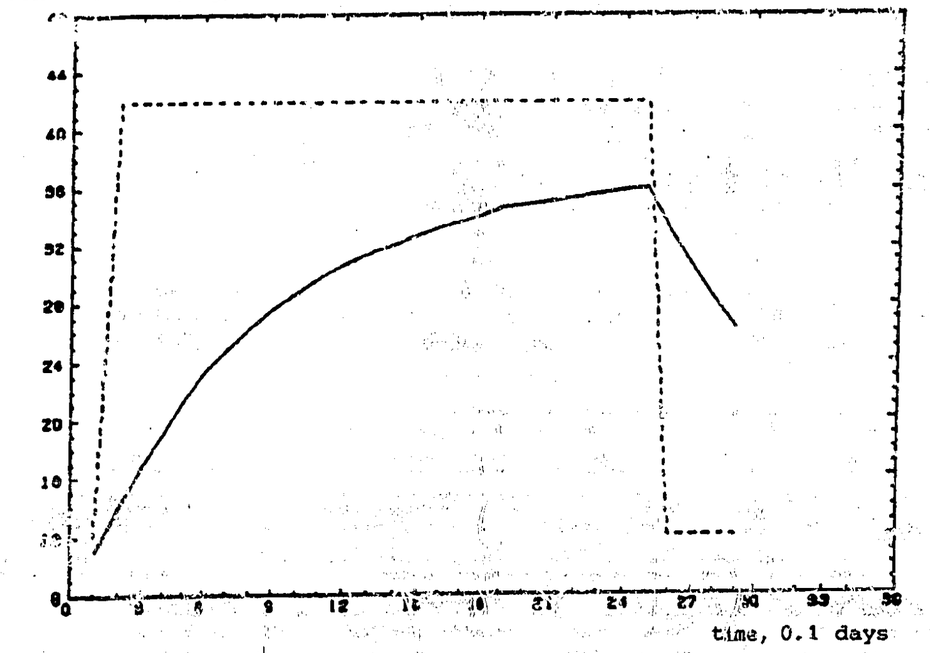


Figure 5.4 Closed and open loop response to change in river nitrate

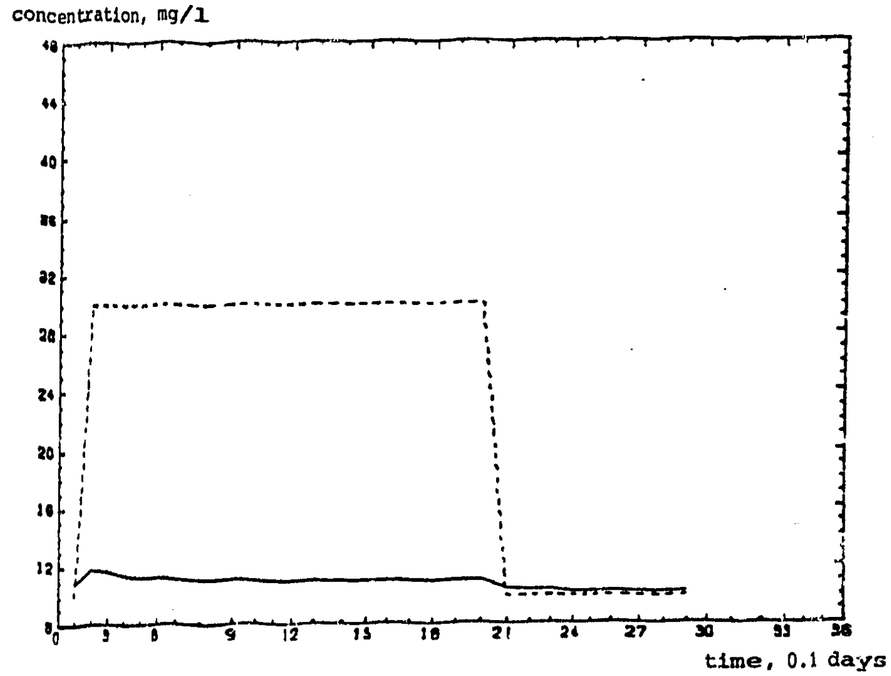


Figure 5.6 Closed and open loop response to change in river nitrate

A further aspect that might be considered is the possibility of using nitrate forecasts to protect the supply against any high nitrate levels which are beyond the blending capacity. This could be achieved by lowering the lagoon nitrate to a level below the standard and so providing some buffering capacity.

Figure 5.6 shows daily LP control in response to a two-day nitrate impulse of 50 mg/l from a base level of 12 mg/l. In the case when $\{n_a = 1, Q_a = 1, n_r = 4, Q_r = 3\}$ the maximum nitrate concentration that can be assimilated without exeedance of the standard is 42 mg/l.

It can be seen from the lower graph of the figure that daily adjustment of the abstraction rates in response to predicted/known changes in river nitrate is acceptable. The upper graph of the same figure shows the effect of acting one day in advance of the impulse by reducing the standard from 11 mg/l to 5 mg/l. The LP then draws down the lagoon nitrate prior to the pulse. The effect is minimal and merely results in expensive and unnecessary pumping from the aquifer and surface reservoir.

5.4.7 Effectiveness of different methods of nitrate removal

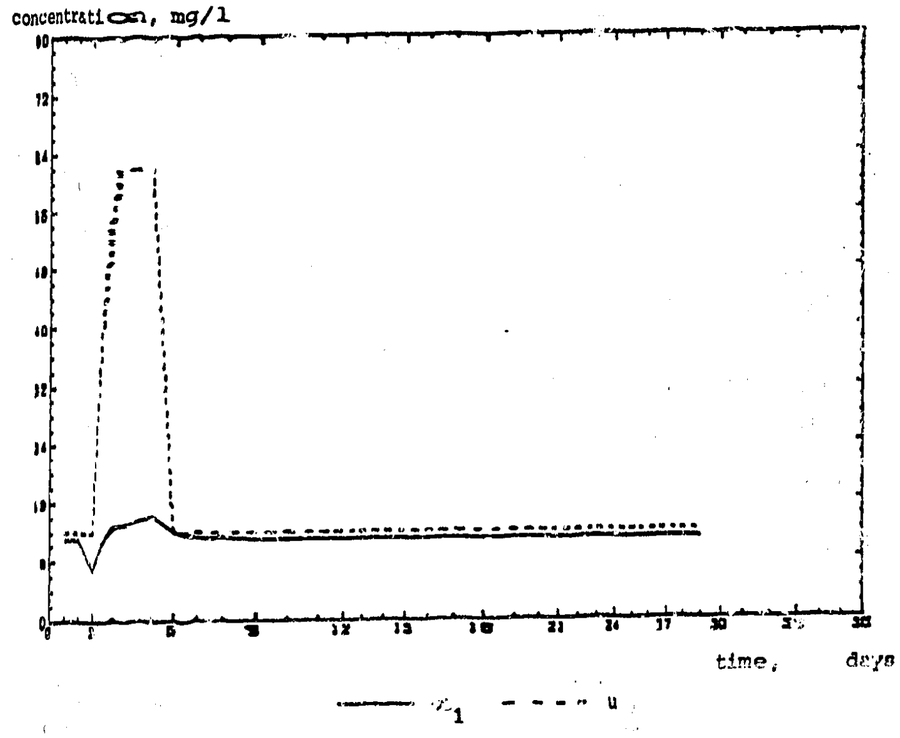
Table 5.6 and Figure 5.7 give some idea of the costs of nitrate removal for a variety of physical and biological methods. The data stem from the 7th Report of the Royal Commission on Agriculture and Pollution (1979) and a report by the Water Research Centre (Gauntlett and Craft, 1979). It should be noted that the cited costs are in terms of nitrate removal per unit volume of water supplied (no explicit indication of the level of nitrate removal is given but a figure of 10 mg/l is typical). A Thames Water Authority Report by Thomas and Smith (1978) gives results of a biological denitrification experiment undertaken at the Lee Bridge Works; this cites a methanol cost of £3.26/Ml for a nitrate removal of 10-15 mg/l. The annual operating cost for this level of removal at a supply of 1 Mgd then compares with that given in Figure

The costs of blending given in both Table 5.6 and Figure 5.9 appear to compare unfavourably with biological denitrification. What is not clear, though, is the purity of the blended sources.

Weighted against the apparent economic advantages of methanol treatment is the speed with which it can be brought on-line to treat shock loads. The Thames report indicates periods of up to 35 days to achieve a fully active biomass giving 90% efficiency.

Clearly, the advantages of the alternatives need to be assessed in terms of the duration and magnitude of adverse nitrate levels. For example the improvement in nitrate concentration in supply resulting from continuous nitrate monitoring and feedback control is self evident. The factors that need to be balanced against this quality improvement are the costs of providing 'continuous' nitrate monitoring in the lagoon, the cost of variable speed pumps/variable sluice gate control and the cost of the lagoon.

For the periodic LP control some less frequent (say daily) measurements of in-stream nitrate concentrations are required; these often form part of any sound river monitoring procedure. Periodic LP control is likely to offer a more cost-effective form of on-line blending than continuous state variable feedback control, even though the latter is the only sure way of dealing with undetected or unpredicted short term transients.



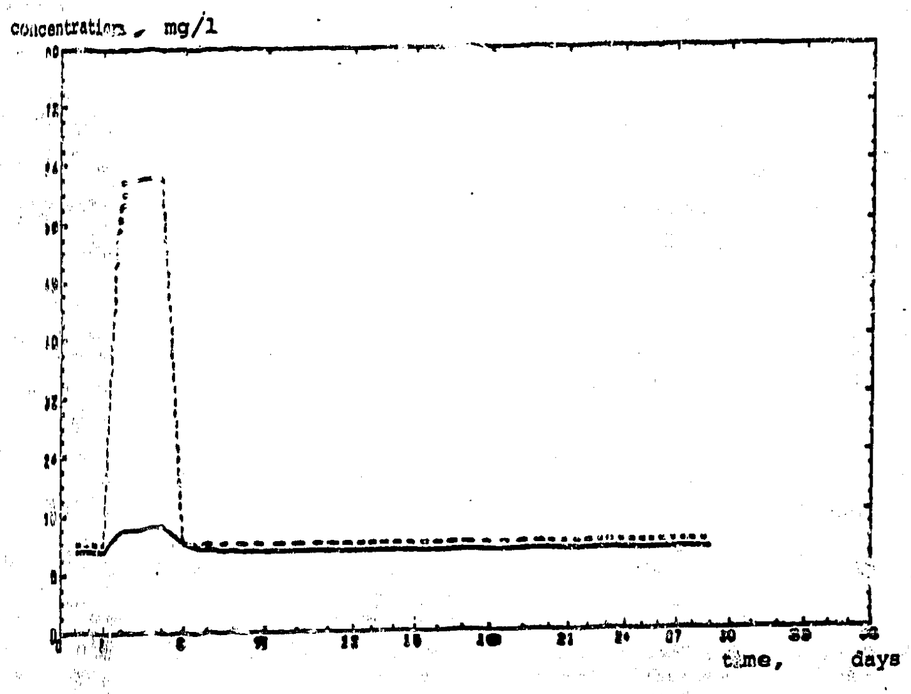


Figure 5.6 Daily LP control. Upper figure has
$x_1^* = 11$ mg/l for $t < 2$, $t \geq 3$
$x_1^* = 5$ mg/l for $t = 3$

	pence per cubic metre supplied
ion exchange	2.0 to 2.5
biological denitrification	0.8 to 1.0
reverse osmosis	5.0 to 6.0
storage - 5,000 cu m for 28 days	1.7
for 6 months	6.0
500,000 cu m for 28 days	0.4
for 6 months	1.4
blending	3.8 to 9.0

Table 5.3 Costs of nitrate removal per cubic metre of water supplied

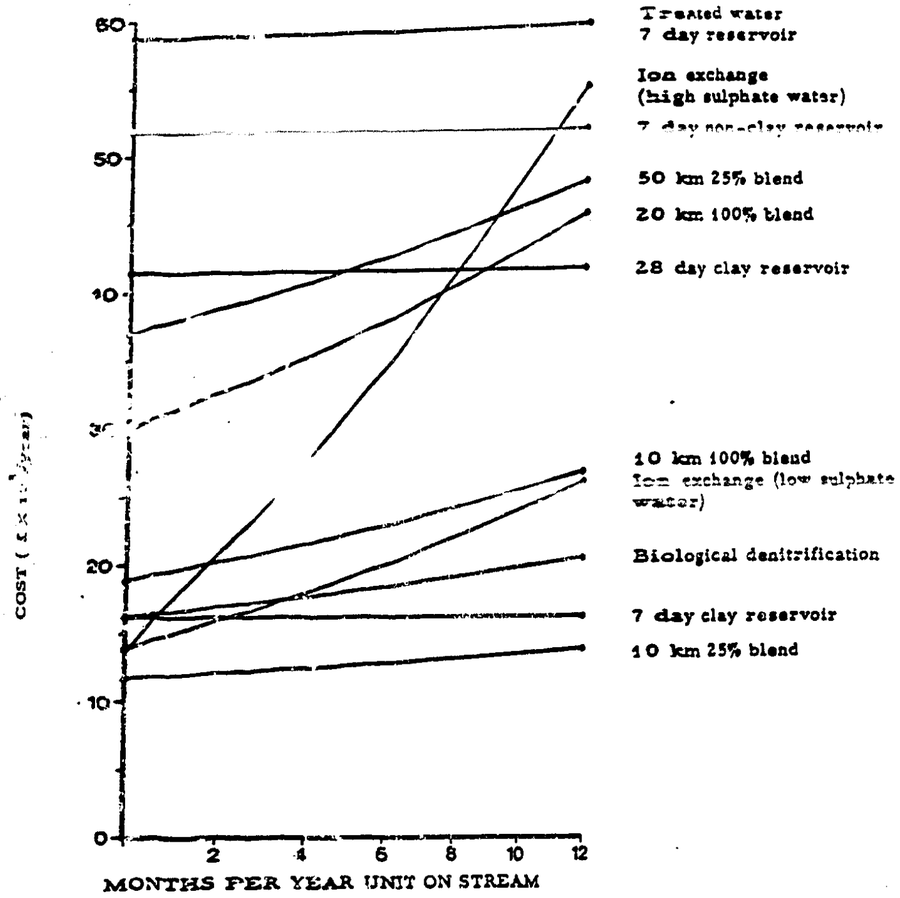


Figure 5.7 Comparative costs of alternative methods for nitrate reduction for 1 mgd (Autumn 1976)

REFERENCES

Andrews, J.F. (1974). Dynamic models and strategies for wastewater treatment processes. Wat. Res. 8: 261-289.

Auslander, D.M., Spear, R.C. and Young, G.F. (1981): Design of control systems with Uncertain Parameters. Journal of Dynamic Systems, Measurement and Control. In Press.

Beck, B. (1978). A Comparative Case Study of Dynamic Models for BO-BOD-Algal Interactions in a Freshwater River RR-78-19. International Inst. for Applied Systems Analysis.

Beck, M.B.(1979). System Identification, estimation and forecasting of water quality, part 1: theory Working paper - 79-31. International Institute for Applied Systems Analysis, Austria.

Beck, M.B. (1979). On line estimation of nitrification dynamics. IIASA Prof. Pap. PP-79-3, IIASA, Laxenburg, Austria.

Beck, M.B. (1981). Operational Water Quality Management: Beyond Planning and Design , Executive Report 7, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Beck, M.B., Latten, A., and R.M. Tong. (1978). Modelling and operational control of the activated sludge process in wastewater treatment. IIASA Prof. Paper. PP-78-10, IIASA, Laxenburg, Austria.

Beck, M.B. and Young, P.C. (1976). Systematic Identification of DO-BOD model structure. Am. Soc. Civ. Engrs. Envir. Engng. Div. 102, 909-924.

Bedford Ouse Study - see Whitehead et al. (1979, 1981).

Bowles, B.A. (1978). Phytoplankton Populations of the River Thames, PhD Thesis, London University.

Busby, J.B. and J. F. Andrews. (1975). Dynamic modeling and control strategies for the activated sludge process. JWOCF 47: 1055-1080.

Cook, S.C. (1981). Use of the extended Kalman filter in the identification of Oxidation ditch dynamics. Presented at the International Workshop on Waste Water Treatment process control, Florence, Italy. Viene 1981.

Davis, A.W. (1977). Pollution problems arising from the 1975-76 drought. In: Scientific Aspects of the 1975-76 Drought in England and Wales. The Royal Society.

Delaware Study - see Thomann (1972)

Fritsch, F.E. (1902). Preliminary report on the Phytoplankton of the Thames, Ann. Bot. 16 (43), 1-9.

Fritsch, F.E. (1903). Further observations of the Phytoplankton of the River Thames. Ann. Bot. 17 (48), 631-647.

Fritsch, F.E. (1905). The plankton of some English rivers. Ann. Bot. 19, 163-167.

Fukunaga, K. and W. L. G. Kootz,(1970). Application of the Karhunen-Loeve expansion to feature selection and ordering. IEEE Trans. on computers. C-19: 311-318.

Gauntlett, R.B. and Craft, D.G. (1979). Biological removal of nitrate from river water. Water Research Centre Technical Report TR90.

Green, L.A. (1978). Nitrates in water supply abstraction in the Anglian region: current trends and remedies under investigation. Water Pollution Control, 478-491.

Hodgkin, E.P., Birch, P.B., Black, R.E. and Humphries, R.B.: (1980). The Peel-Harvey Estuarine System Study. Report No. 9, Department of Conservation and Environment, Western Australia.

Hornberger, G.M. and R. C. Spear,(1980). Eutrophication in Peel Inlet: I. The Problem defining behaviour and a mathematical model for the phosphorous scenario. Water Research 14: 29-42.

Hornberger, G.M. and R. C. Spear,(1981). An approach to the preliminary analysis of environmental systems. J. of Environ. Mgmt., 12, 7-18.

Humphries, R.B., Beer, T. and Young, P.C. (1980). Weed Management in the Peel Inlet of Western Australia. Water and Related Land Resource Systems (Ed: Y.Haimes and J. Kindler), IFAC, 131-139.

Jamieson, D.G. (1980). Management of Water Quantity and Water Quality: A Future Scenario. Paper to IH/WRC Conference, Institute of Hydrology.

Jazwinski, A.H., (1970). Stochastic processes and filtering theory. Academic Press, New York.

Kendall, M.G. and A. Stuart, (1969). The Advanced Theory of Statistics, Vol. 3, Hafner Pub. Co.

Kneese, A.V. (1962). Water Pollution: Economic Aspects and Research Needs. Resources for the Future Report.

Kowalczewski, H. and Lack, T.J. (1971). Preliminary production and respiration of the Phytoplankton of the River Thames and Kennet at Reading. Freshwater Biol. 1, 197-212.

Kynch, G.J. (1952). A theory of sedimentation. Trans. Faraday Soc. 48: 166-176.

Lack, T.J. (1971). Quantitative studies on the Phytoplankton of the River Thames and Kennet at Reading, Freshwater Biol., Vol. 1, 213-224.

Lederman, T.D., Hornberger, G.M. and Kelly, M.G. (1976). The calibration of a Phytoplankton growth model using batch culture data, University of Virginia, Dept. of Environmental Science Report.

Ljung, L. (1979). Assymptotic behaviour of the extended Kalman filter as a parameter estimator for linear systems. IEEE Trans. Automatic Control, AC-24(1), pp. 36-50.

Lund, J.W.G. (1949). Studies of Asterionella (1). Origin and nature of cells producing seasonal maxima. J. Ecol. 37, 389-419.

Marsili-Libelli, S. (1980). Reduced order modeling of activated sludge process. Ecol. Mod. 9: 15-32.

Marsili-Libelli, S. (1980). Optimal aeration control for wastewater treatment. In: Cuenod, M.A. (ed.) (IFAC) Computer Aided Design of Control Systems, Pergamon Press, New York, pp. 511-516.

Olsson, G. and J.F. Andrews. (1978). The dissolved oxygen profile - a valuable tool for control of the activated sludge process. Wat. Res. 12: 985-1004.

O'Connell, P.E. (1981). (Editor) Proceedings of First International Workshop, Institute of Hydrology Report.

O'Connor, D.J., Thomann, R.V. and Di Toro, C.M. (1976) Ecologic models. In: Systems Approach to Water Management, McGraw-Hill, New York.

Onstad, C.A. and Blake, J. (1980). Thames nitrate and agricultural relations. Proc. ASCE Watershed Management Symposium, 961-973, Boise, I.D.

Owens, M., Garland, J., Hart, I.C. and Wood, G. (1972). Nutrient budgets in rivers. Symp. Zool. Soc. Land., No. 29, 21-40.

Rice, C.H. (1938). Studies in the Phytoplankton of the River Thames 1 and 2. Ann. Bot. 2, 539-557; 559-581.

Smeers, V. (1980). On the Economics of Time Varying River Quality Control Systems, Proceedings of Real Time Water Quality Management, (Editor M Beck) IIASA Report CP-80-38.

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

Spear, R.C. and G.M. Hornberger. (1981). Solar power satellite: energy analysis revisited. Space Solar Power Reviews, In press.

Spofford, W.O., C.S. Russel and R.A. Kelly (1976). Environmental Quality Management - An Application to the Lower Delaware Valley. Resources for the Future Research Paper R-1, Washington D.C.

Stenstrom, M.K. and J.F. Andrews. (1979). Real time control of the activated sludge process. ASCE J. Env. Eng. Div. 105: 245-260.

Swale, E.M.F. (1962). Studies in the Phytoplankton of a calcareous river. PhD Thesis, University of London.

Tanuma, M. (1980). Water quality management in a wastewater treatment plant. In: Beck, M.B. (ed.) Real time water quality management,

Thomann, R.V. (1972). Systems analysis and water quality management, McGraw-Hill, New York.

Thomas, E.V. and Smith, A.P. (1978). Biological denitrification of river derived water; experimental filters at Lee Bridge Works. Thames Water Authority PU140 (presented at IWES AGM (SE Section), April 1978).

Toms, I.P., Mindenhall, M.J. and Harman, M.M.I. (1975). Factors affecting the removal of nitrate by sediments from rivers, lagoons and lakes RC Technical Report, TR14.

Tong, R.M. ( ). A Control Engineering Review of Fuzzy Systems, Automatica 13, pp. 559-569.

Trent Model Study (1968). Institution of Water Engineers Report.

Tyteca, D., Smeers, Y and E. J. Nyns. (1977). Mathematical modeling and economic optimization of wastewater treatment plants. CRC Critical Reviews in Environmental Control 8: 1-89.

Venn, M.W. and Day, B. (1977). Computer aided procedure for time series analysis and identification of noisy processes (CAPTAIN). Inst. of Hydr. Report No. 39, Wallingford.

Whitehead, P.F. (1979). Applications of recursive estimation techniques to time variable hydrological systems. J. Hydrol. 40, 1-16.

Whitehead, P.G. (1981). An instrumental variable method of estimating differential equations of dispersion and water quality in tidal rivers. J. Ecol. Model., 9, 1-14.

Whitehead, P.G., Young, P.C. and Hornberger, G. (1979). A systems model of flow and water quality in the Bedford Ouse river system. I - Streamflow modelling. Water Res. 13, 1155-1168.

Whitehead, P.G., Beck, M.B. and O'Connel, P.E. (1981). A systems model of flow and water quality in the Bedford Ouse River System. II - Water Quality Modelling. Wat. Res., 15, 1157-1171.

Whitehead, P.G. and Hornberger, G.B. (1984). Modelling algal behaviour in river systems. Water Research (in press).

Whitehead, P.G. and Young, P.C. (19). Water quality in River Systems: Monte Carlo Analysis. Wat. Resources Res. 15, 451-459.

Whitehead, P.G., Caddy, D.E. and Templeman, R.F. (1983). An on-line Monitoring, Data Management and Forecasting System for the Bedford Ouse River. IAWPRC Third RCS River Basin Management Conference, York.

Whitehead, P.G. and Williams, R.J. (1984). Determining Dispersion Coefficients from Tracer Data. In preparation.

Young, P.C. (1978). A general theory of modeling for badly defined systems. In: Vansteenkiste, G.C. (ed.). Modeling, Identification and Control in Environmental Systems, North Holland Pub. Co., Amsterdam, pp. 102-135.

Young, P.C. and Beck, N. (1974): The Modelling and Control of Water Quality in a River System, Automatica 10: 455-468.

Appendix A: The Bedford Ouse Monitoring, Modelling and Forecasting Scheme

A.1 Background

The Water Industry is often sceptical about the value of both continuous water quality monitoring and applying mathematical models of water quality to river systems. Considerable benefits can accrue, however, from combining these approaches by telemetering flow and quality data to a computer and using a model to forecast likely changes in flow and quality with time. In a collaborative study between the Anglian Water Authority and the Institute of Hydrology the aim has been to forge a new tool which can be of value to personnel concerned with river pollution control and with the responsibility for forecasting water quality changes at important abstraction points.

To make good operational decisions relating to river quality, there is a requirement for real time data. For example, when evaluating the risks to river conditions associated with a loss of efficiency at an effluent treatment plant it is necessary to estimate the resulting short term changes in river water quality. A subjective approach can be taken where an inspector draws on many years of local experience to assess a pollution event. An alternative approach is to use a computerised mathematical model capable of forecasting flow and water quality along a river system. The particular approach used must enable action to be taken in time to anticipate any critical situations. For example, it may be necessary to close a direct abstraction to a water supply treatment plant whilst a pulse of polluted water flows past. By having prior warning of the time of arrival and the severity of the polluting conditions, it is possible to increase operational efficiency as well as safeguard water supply.

A.2 Monitoring equipment

Reference to Figure A.1 shows the extent of the Great Ouse network on the Bedford Ouse; another three stations not shown are located on the Ely Ouse river system, whilst a further two stations are planned, as shown, at Olney and Foxcote. The monitoring stations and flow gauges are linked to a central microcomputer at Cambridge by a telemetry system shown schematically in Figure A.2. Details of the hardware used are given by Whitehead, et al., 1983. The decisions as to where to develop stations were largely determined by the findings of the Bedford Ouse study (1975) which was established in 1972 by the Great Ouse River Authority and the Department of Environment. In addition to investigating long term planning problems in the Bedford Ouse area, a critical reach of the river between the New City of Milton Keynes and the water abstraction intake at Bedford was studied to assess the likely impact of effluent on water quality. In particular, with limited bankside storage available at Bedford for public water supply, there was concern that a significant pollution would affect plant operation. In order to investigate the operational problems that could arise, a mathematical model was developed for this reach of this river based on daily or more frequent data collected using continuous monitors (Whitehead et al 1979, 81). Since 1975 the Bedford Ouse monitoring stations have been installed with the aim of monitoring the

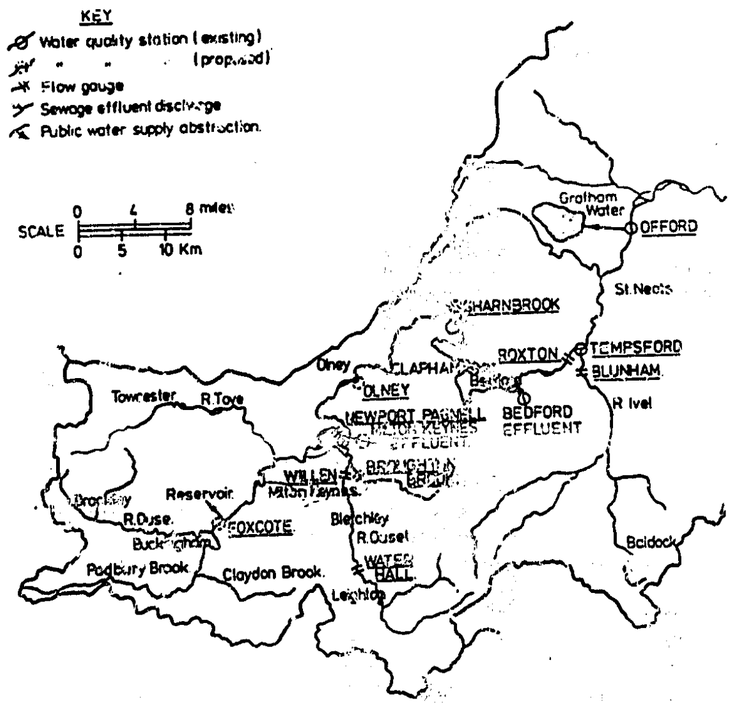


Figure A.1. Bedford Ouse System showing the Deployment of Water Quality Monitors and Flow Gauging Stations

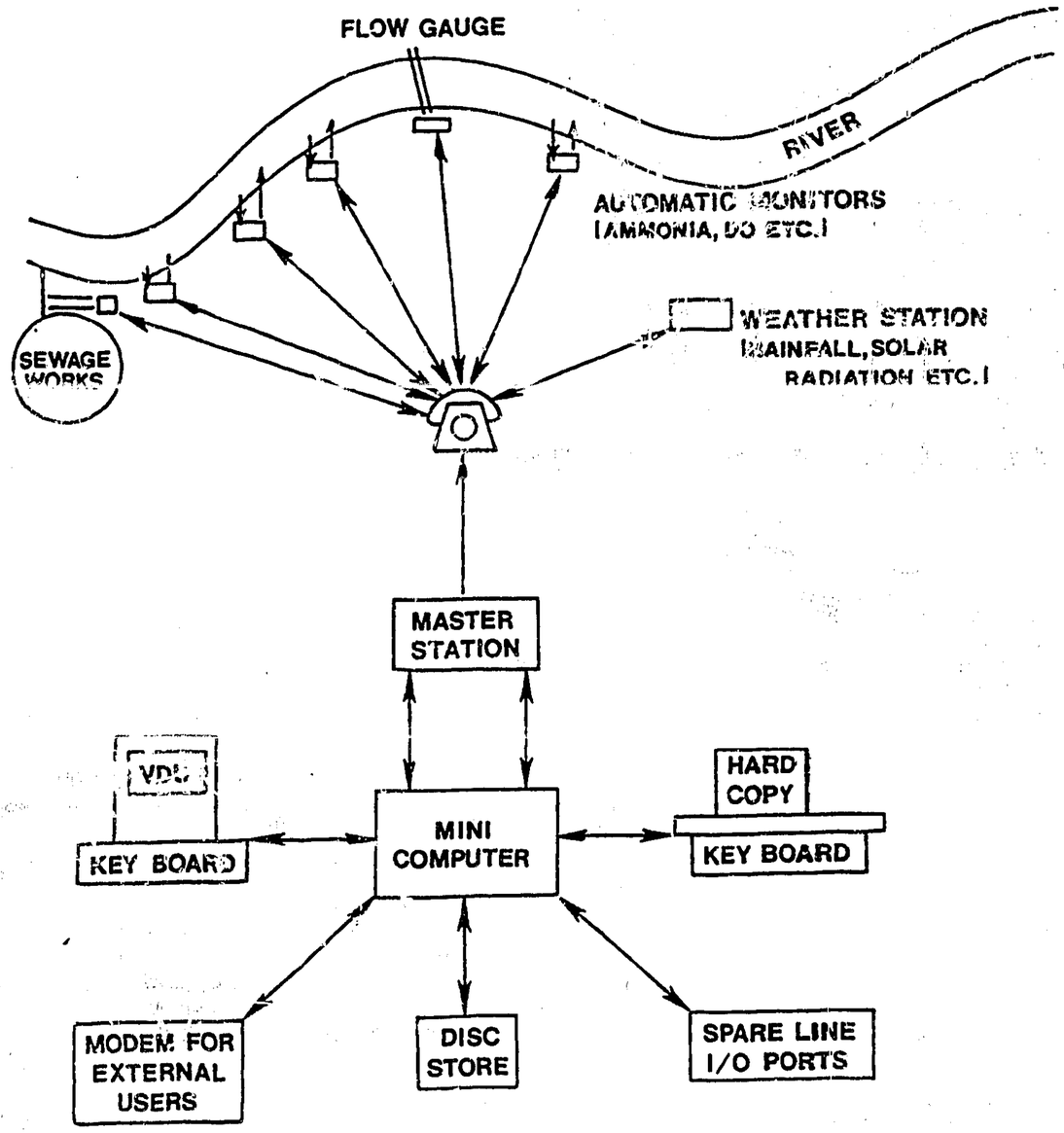


Figure A.2 Water Quality Monitoring, Telemetry and Mini-Computer Processing System for the Bedford Ouse.

impact of effluent discharges and protecting river abstraction sources for potable supply. In choosing which variables to measure, the philosophy has been to use only sensors which are reliable for long-term duty. Variables such as ammonia, dissolved oxygen, conductivity, pH, temperature and nitrate have been monitored on the Bedford Ouse system.

A.3 Data acquisition and processing

Data are relayed to the centre at Cambridge where equipment is situated for telemetry control, data processing and report production. The scanning of the flow gauges and water quality monitors is computer controlled with the outstation telephone numbers, which are held in computer memory, being fed to the master station (autodialler/receiver). After making contact with an outstation, data are transmitted and checked before being stored in the computer memory for further processing. Information can be presented to operating staff in several forms and data summaries, alarm messages and graphical displays are available on either the printer or on the visual display unit. The computer system can be operated by staff not qualified in computer science and the Institute of Hydrology has devised an interactive system which prompts the user into selecting options which appear on the visual display unit.

An important function of the micro-computer is the storage of data and this is achieved using two double density floppy discs capable of storing several weeks' data for all the monitoring stations. Data stored on the floppy discs can be retrieved at any time for direct listing or for the calculation of summary statistics such as daily mean, maximum, minimum and 95 percentiles. The logging program on the micro-computer allows the acquisition of both analogue data (ie measurements from water quality instruments) and digital signals related to equipment status (eg pump failure). Comparison of data with preset alarm levels enables warning messages to be given in the event of a pollution incident. Given such warnings the situation can be investigated using the mathematical model to obtain forecasts of flow and water quality up to 80 hours ahead, given certain assumptions on the upstream conditions during the forecast period.

The options available on the micro-computer system are selected using a simple interactive system by which the operator is prompted to answer questions displayed on the screen. The options available may be summarised as follows:-

(a) *Start Logging*, which allows the operator to start the collection and storage of data from the outstations at predetermined times. The logging program also converts the instrument signals to river flow or concentration units, allows for calibration factors and prints out alarm level messages, equipment status and daily statistics;

(b) *Interrogate Outstation* which allows the operator to interrogate any outstation of his/her choice immediately. This is particularly useful if a pollution incident has occurred, or when installing and maintaining instruments;

(c) *Plot Data* which provides the operator with a choice of various graphical presentations of data on the visual display screen. Plots may be obtained for the data collected at any outstation, see Figure A.3.

(d) *Edit Master File* which allows the operator to create or edit the masterfile which holds all the fixed parameters describing each outstation's configuration and other attributes such as site name, telephone number, type of instruments, equations defining data conversions, calibration factors, alarm levels and warning messages and equipment reliability;

(e) *Print Master File* which allows the operator to list the contents of the master file on the DEC printer;

(f) *Print dial out statistics* which reports the number of successful and failed dial out attempts for each outstation;

(g) *Initialise disc* which allows a floppy disc to be prepared for use on the micro-computer system;

(h) *Create new system* allows the operator to copy a system disc and prepare it for use;

(i) *Initialise data files* sets up a data file for each outstation;

(j) *Delete files* deletes the current data files for each outstation;

(k) *Print daily, weekly, monthly and annual summaries* prints either a daily weekly, monthly or annual summary (mean, maximum, minimum, standard deviation, 95 percentile) on the DEC printer;

(l) *Run flow and water quality model* allows the system to run the flow and water quality model using data [illegible];

(m) *Run impulse model* asks the system to run the impulse model using data supplied by the operator on the [illegible] of pollutant (BOD, ammonia or conservative pollutant), pollutant [illegible] rate, concentration and location of discharge;

(n) *List station data* prints the output data for a specified station;

The programs for all these options have been written in standard FORTRAN so that the system can be easily modified as Water Authority requirements change.

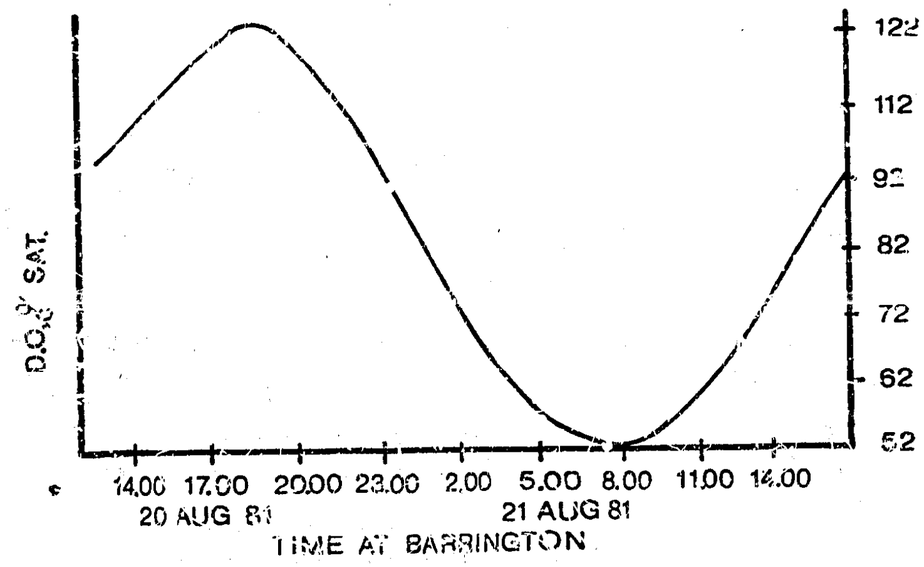


Figure A.3. Dissolved Oxygen monitored at Barrington on the River Cam

A.4 Modelling Flow and Water Quality in the Bedford Ouse

The model developed for the Ouse is an operational model and is therefore designed to use time varying input (upstream) measures of flow and water quality to compute time varying output (downstream) responses. The model characterises the short term (hourly) system behaviour and provides a mathematical approximation to the physico-chemical changes occurring in the river system. The structure of the model developed for the Bedford Ouse is shown in Figure A4; a multi-reach flow model is linked with the water quality model so that flow/quality interactions are incorporated directly.

A detailed description of the flow and quality model is described in Whitehead *et al* (1979, 81); a summary is included here. The model is based on a twenty reach representation of the Bedford Ouse River, in which each reach is characterised by a number of compartments. The model for flow variations in each compartment is based on an analogy with the mass balance equations for the variations in concentration of a conservative pollutant under the assumption of uniform mixing over the compartment. The streamflow model is similar to the Thames model described by equations (4.8) - (4.11) in Section 4.2.2.4 of the proceedings.

In order to evaluate the Bedford Ouse velocity-flow relationships, a series of tracer experiments has been conducted on the river. A known mass of iodide was injected into the river and the iodide concentration is determined at one or more selected locations downstream either continuously using selective ion detection equipment or by sampling the river water and subsequent analysis of the samples at the Institute of Hydrology. Using information on velocity and flow rate from these experiments and on earlier experiments conducted in 1975, the parameters in equation (4.10) have been determined as follows:

$$a = 0.045 \text{ and } b = 0.67$$

Given information on upstream and tributary inputs, the flow routing model simulates stream flow by solving equation (4.11) with T defined through equations (4.7) and (4.10) as in the case of the Thames model.

The water quality models for the Bedford Ouse are similarly based on a mass balance principle but include factors to allow for the non-conservative nature of water quality variables. For example dissolved oxygen in the river is a balance between the various sources and sinks of oxygen. On the one hand there is oxygen supplied by the reaeration from the atmosphere and photosynthetic oxygen produced by plants and algae and, on the other hand, oxygen is being consumed by respiration processes and the removal of oxygen during the bacterial breakdown of organic materials from effluents. The mass balance equations developed to simulate water quality behaviour are as follows:

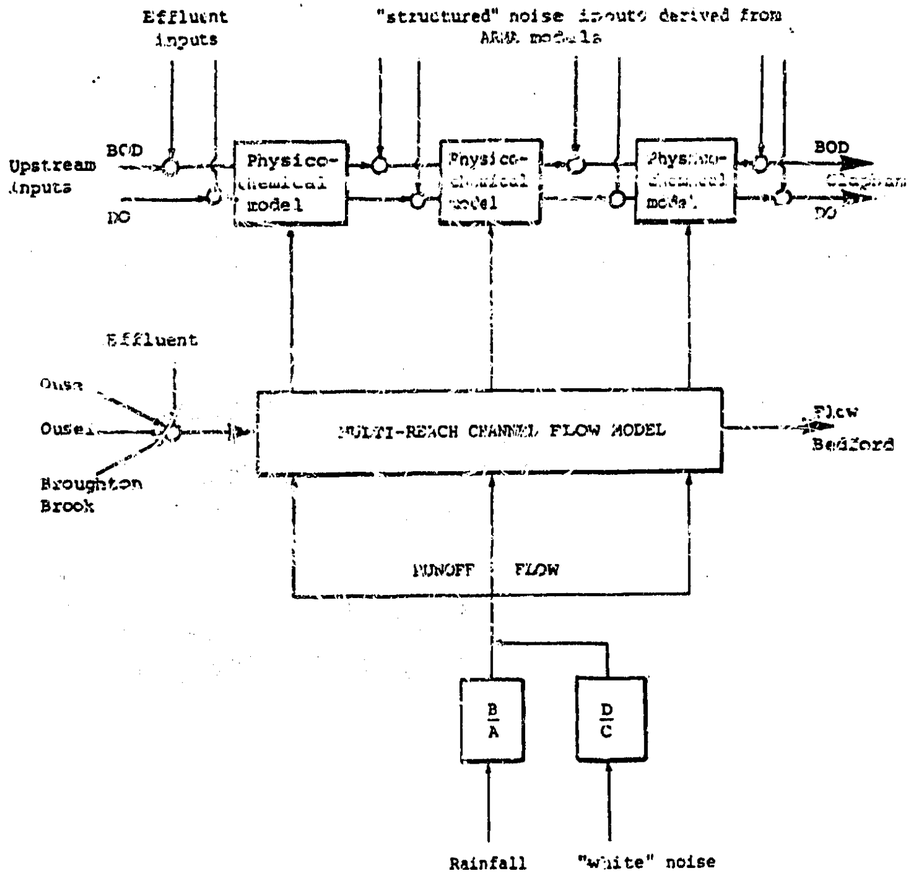


Figure A.4. Bedford Ouse Model Structure

Chloride

$$\frac{dx(t)}{dt} = \frac{Q_i(t)}{V} u_i^1(t) - \frac{Q_o(t)}{V} x_1(t) + S_1(t) \qquad (A.1)$$

Nitrate

$$\frac{dx(t)}{dt} = \frac{Q_i(t)}{V} u_i^2(t) - \frac{Q_o(t)}{V} x_2(t) - k_1 x_2(t) + S_2(t) \qquad (A.2)$$

Ammonia

$$\frac{dx_3(t)}{dt} = \frac{Q_i(t)}{V} u_i^3(t) - \frac{Q_o(t)}{V} x_3(t) - k_2 \left(\frac{1}{Q_o(t)}\right) x_3(t) + S_3(t) \qquad (A.3)$$

Dissolved Oxygen (DO)

$$\frac{dx_4(t)}{dt} = \frac{Q_i(t)}{V} u_i^4(t) - \frac{Q_o(t)}{V} x_4(t) - 4.33\ k_2 \left(\frac{1}{Q_o(t)}\right) x_3(t) - k_3 x_5(t)$$

$$+ k_4(C_s(t) - x_4(t)) + S_4(t) \qquad (A.4)$$

$$\frac{dx_5(t)}{dt} = \frac{Q_i(t)}{V} u_i^5(t) - \frac{Q_o(t)}{V} x_5(t) - k_3 x_5(t) + S_5(t) \qquad (A.5)$$

where x refers to the downstream (reach output) concentration mg/l;
u refers to the upstream (reach input) concentration mg/l;
Q is the flow rate (determined from the flow model) and V is the reach volume;
S refers to the additional sources and sinks affecting water quality such as the net rate of addition of DO in the reach by photosynthetic/respiration activity of plants
C_s is the saturation concentration of dissolved oxygen.

The model has been programmed to run in two modes.

(1) normal operating mode, in which information on upstream input flows and quality are taken from the outstations and forecasts produced for all downstream reach boundaries up to 80 hours ahead. Because of the considerable travel time on the river, downstream forecasts are based on measured upstream flow and quality. It is necessary, however, to forecast upstream conditions to provide reasonable forecasts for the upper reaches. Techniques for this are given in Whitehead *et al* (1979).

(2) Impulse mode, in which the operator can supply information on an upstream impulse discharge of a conservative pollutant, ammonia or BOD. The resultant simulation uses the current flow and river quality data and simulates the slug of pollutant moving down the river system. Again forecasts up to eighty hours ahead are available at all of the twenty reach boundaries.

It is assumed that weather conditions remain stable during this forecast period. It would be possible however to link existing rainfall gauges into the telemetry scheme to forecast flow and quality.

Figure A5 and A6 show a typical simulation in impulse mode. The output data from the model can be plotted either as a profile down the river at a specific time or else as a function of time at any selected reach boundary. Thus if information on say the location of the minimum is required then a river profile would be plotted. If, however, it is required to know the likely time of arrival of a slug of pollutant at a given point the time-concentration curves would be plotted.

A.5 Case Studies

(a) Milton Keynes Effluent

The operational aspects of the model have already been tested in a real situation when Bedford Sewage Division reported a loss of oxidative treatment at the Cotton Valley Sewage Treatment works and effluent containing elevated levels of ammonia was discharged into the river at Newport Pagnell. In order to protect the water supply at Bedford (Clapham abstraction) information was required on the likely ammonia concentrations at Clapham and the time of arrival of the pollutant. Simulating ammonia using the model indicated an arrival time of four days with concentrations of ammonia of 1.2 mg/l at Clapham. The observed levels at Clapham were 1.12 mg/l and the arrival time 4 days 2 hours. In this situation the model provided valuable information for the management at the Clapham water abstraction plant and gave them effectively a four day warning.

(b) Bedford Effluent

The second case study relates to the release of unsatisfactory effluent from Bedford sewage works which resulted in a significant pulse of ammonia being discharged to the river. Again running the model in an impulse mode gave reasonable forecasts of ammonia and dissolved oxygen concentrations downstream at Tempsford (see Figure A7)

In both these situations information of value to pollution inspectors has been provided using the computer programs, monitoring and telemetry system. Such forecasts are of operational use and the flexible computing system installed on the Bedford Ouse is essential if full use is to be made of the collected data.

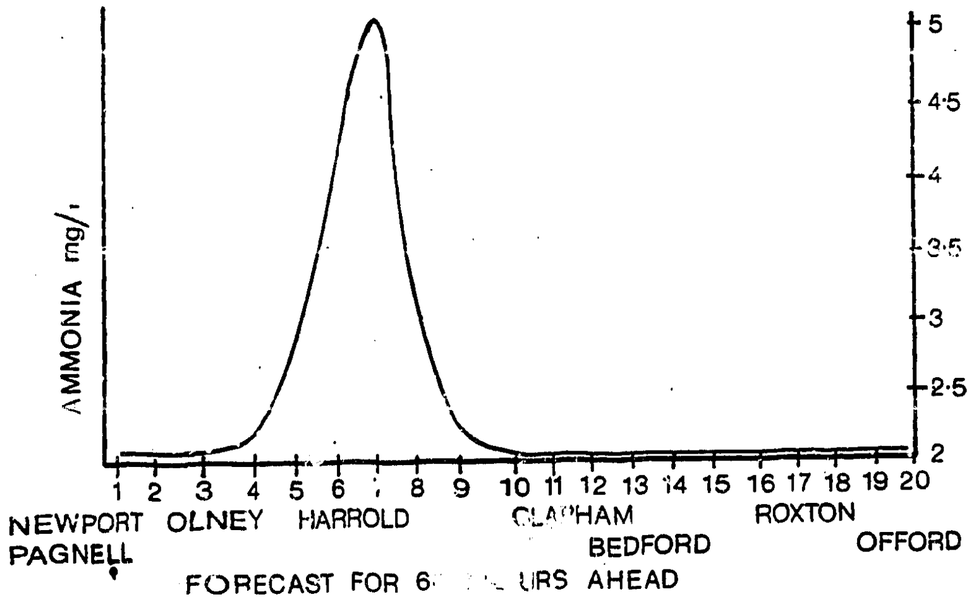


Figure A.5. Impulse Profile along River

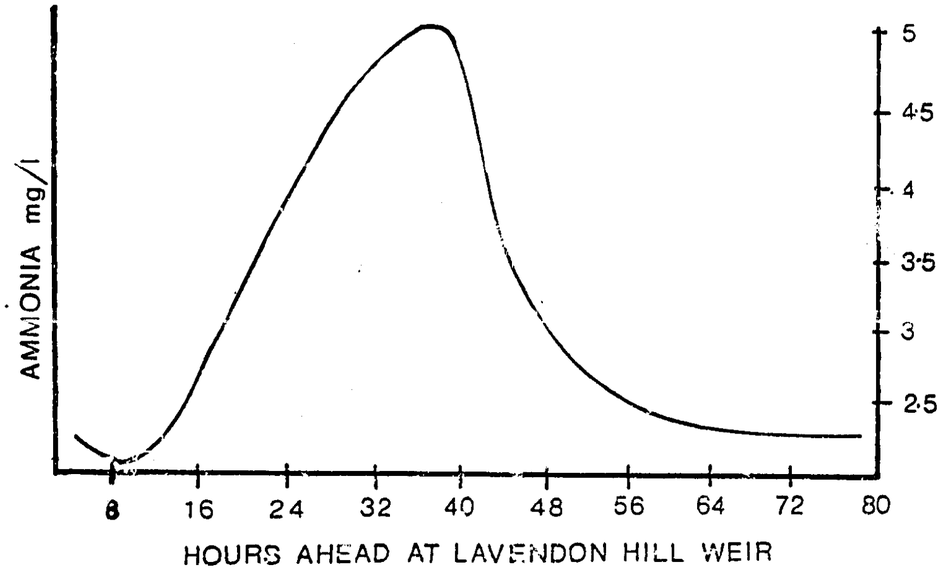
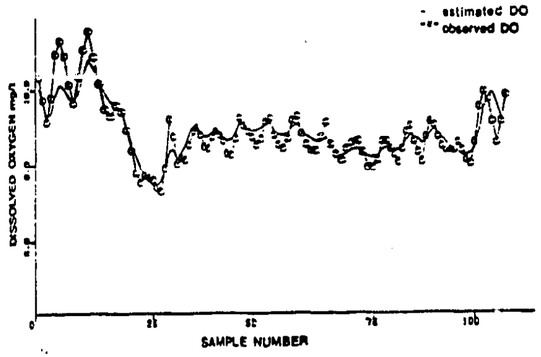


Figure A.6. Impulse concentration against time.



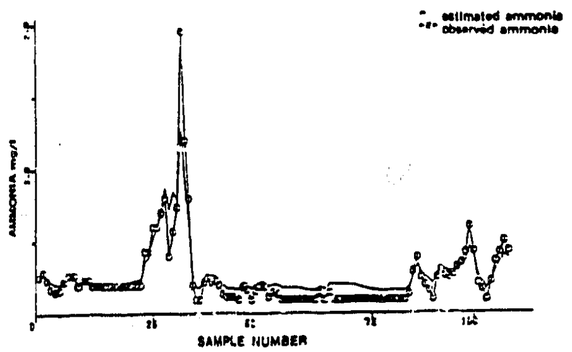


Figure A.7. Ammonia and DO forecasts at Tempsford

A.6 Conclusions

Mathematical models have been developed to provide long term predictions and short term forecasts of water quality changes in river systems. The models are dynamic so that transient violations of water quality standards can be investigated. Information on the extremes of river behaviour are often required to assess the reliability of a particular system design and in the case of the Thames Study the model will be used to investigate alternative strategies for managing nitrate. In the case of the Bedford Ouse, the model is linked to a continuous water quality monitoring scheme and short term forecasts are available for operational management.

Appendix B: Computational aspects of the Extended Kalman Filter (EKF), with user guide to computor prograr

B.1 Computational aspects

A computer program for the solution of the non-linear estimation problem by the EKF has been written and applied in several modelling problems. The program has been developed for continuous/discrete systems and is presented together with the input format and operating instructions.

Several computational problems have been encountered in the modelling studies and it will be useful to consider these here. Recall the definition of the transition matrix in Section 3.3

$$\phi(t_k, t_{k+1}, \hat{\underline{x}}(t_k|t_k)) = \exp\{\int_{t_k}^{t_{k+1}} F'(\hat{x}(\tau|t_k), \underline{u}(\tau), \tau) d\tau\} \qquad B.1$$

For computational simplicity consider

$$\phi(t_k, t_{k+1}, \hat{\underline{x}}(t_k|t_k)) = \exp\{F(\hat{\underline{x}}(t_k|t_k), \underline{u}(t_k), t_k)(t_{k+1}-t_k)\} \qquad B.2$$

as an approximation to B.1. Consider also the matrix exponential defined by

$$\exp\{\underline{A}\Delta t\} = I + \underline{A}\Delta t + \underline{A}^2\frac{\Delta t^2}{2!} + \ldots + \underline{A}^n\frac{\Delta t^n}{n!} \quad .$$

The series is truncated when

$$A^n(i,j)\frac{\Delta t^n}{n!} < \varepsilon \quad \forall\, i,j$$

where ε is a limit of accuracy of one order of magnitude less than the desired precision of solution - there are two potential sources of error in the above formulation; (i) in the simplified transition matrix definition and (ii) in the method of matrix exponential evaluation.

Transition Matrix Definition

The Jacobian matrix F measures the sensitivity of the systen dynamics to changes in the states and parameters. For linear systems, the sensitivity depends only upon the time-dependency of the parameter set. For

$$\alpha(t) = f(t)$$

where f(t) is a known function, equation (B.1) may be solved analytically. More often, the parameter set is independent of time and the transition matrix definition becomes simple. For non-linear systems however the situation is not so clear and the implications of equation (B.2) must be

We know that

$$P(t_k|t_k) \triangleq \mathrm{cov}(\tilde{\underline{x}}(t_k|t_k)) = \mathrm{cov}(\tilde{\delta\underline{x}}(t_k|t_k))$$

and $P(t_{k+1}|t_k) \triangleq \mathrm{cov}(\tilde{\underline{x}}(t_{k+1}|t_k)) = \mathrm{cov}(\tilde{\delta\underline{x}}(t_{k+1}|t_k))$

where $\tilde{\delta\underline{x}}(t_k|t_k) = \delta\underline{x}(t_k) - \hat{\delta\underline{x}}(t_k|t_k)$

$$\tilde{\delta x}(t_{k+1}|t_k) = \delta\underline{x}(t_{k+1}) - \hat{\delta\underline{x}}(t_{k+1}|t_k)$$

As a result of the simplifying assumption, defined by equation (B.2), let us suppose that the evaluation of ϕ is inaccurate by a significant amount $\Delta\phi$; then assuming no exogenous inputs and Q = 0, if

$$\hat{\delta\underline{x}}(t_{k+1}|t_k) = \phi(t_k, t_{k+1})\hat{\delta\underline{x}}(t_k|t_k)$$

and, dropping the argument list of ϕ for simplicity,

$$\hat{\delta\underline{x}}(t_{k+1}|t_k) = (\hat{\phi} - \Delta\phi)\hat{\delta\underline{x}}(t_k|t_k)$$

where $\hat{\phi} = \phi + \Delta\phi$,

then

$$\begin{aligned}\mathrm{cov}(\hat{\delta\underline{x}}(t_{k+1}|t_k)) &= E(\hat{\delta\underline{x}}(t_{k+1}|t_k).\hat{\delta\underline{x}}^T(t_{k+1}\ t_k)) \\ &= E(\phi\ \hat{\delta\underline{x}}(t_k|t_k).\hat{\delta\underline{x}}^T(t_k\ t_k)\phi^T) \\ &= \phi P(t_k|t_k)\phi^T\end{aligned}$$

and, computationally,

$$\begin{aligned}\mathrm{cov}(\hat{\delta\underline{x}}(t_{k+1}|t_k)) &= E(\hat{\phi}\ \hat{\delta\underline{x}}(t_k|t_k).\hat{\delta\underline{x}}(t_k|t_k)\hat{\phi}^t) \\ &= E((\phi+\Delta\phi)\hat{\delta x}(t_k|t_k).\hat{\delta\underline{x}}^T(t_k|t_k)(\phi+\Delta\phi)^T) \\ &= (\phi+\Delta\phi)P(t_k|t_k)(\phi+\Delta\phi)^T\end{aligned}$$

Hence given the correct estimate of $P(t_k|t_k)$ the value of $P(t_{k+1}|t_k)$ is corrupted by an amount dependent upon the initial error in ϕ (Cook, 1982).

Transition Matrix Computation

The computation of the transition matrix may proceed in a step-wise manner throughout the integration interval. In the computer program, the computation is linked directly to system function integration. The integration is performed using a variable step length, 4th order Runge-Kutta algorithm routine. The step length varies according to a pre-set accuracy of

estimation, so that it is comparatively small in regions of high system non-linearity but large in linear regions. Suppose the integration within-step, step length is h_i, then the transition matrix evaluation is given by,

$$\phi(t,t+\Delta t) = \exp\{\sum_{i=1}^{n} ((F(t,+h_{i-1}) + F(t,t+h_i))\frac{h_i}{2}\}$$

where $F(t,t+h_{i-1})$ = the Jacobian evaluated at the beginning of any integration within-step, step h_i.

$$= F(\hat{x}(t + \sum_{j=0}^{i-1} h_j\ t),\ t,\ t + \sum_{j=0}^{i-1} h_j)$$

$F(t,t+h_i)$ = the Jacobian evaluated at the end of any integration within-step, step h_i.

$$= F(\hat{x}(t + \sum_{j=1}^{i} h_j\ t),\ t,\ t + \sum_{j=1}^{i} h_j)$$

where n = number of steps taken within the integration period.

This assumes that the non-linearity in F is proportional to the non-linearity in $\hat{x}(.|.)$, which may seem unreasonable. For example if we consider the following non-linear, scalar system,

$$\dot{x}(t) = x(t)^a$$

$$F = \frac{\partial \dot{x}(t)}{\partial x(t)} = ax(t)^{a-1}$$

so that the proportionality is lost. However, the system function/Jacobian evaluation link is a computational convenience which is designed to improve upon any arbitrary computational method.

Matrix Exponential Evaluation

The second problem to arise in the transition matrix computation is that of errors in evaluating the matrix exponential. The method employed has been stated but this is subject to computational difficulties, most notably the occurrence of terms of large magnitude in the series prior to truncation. This can be overcome in a number of ways, including

(i) using the matrix characteristic equation to give an analytical solution and

(ii) employing a matrix factorisation. The latter method is chosen since evaluation of the matrix eigenvalues is avoided. Consider therefore,

e^F where F is an nxn matrix representing the summed Jacobian matrix

$$e^F = (e^{(F/\alpha)})^{\alpha}.$$

If α is a factor of 2 such that

$$\alpha = 2^{\beta}$$

then,

$$e^F = \prod_{i=1}^{\beta} [e^{(F/\alpha)}]^2$$

where α is chosen to be close to the root mean square trace of the matrix F (the r.m.s. is used to avoid computational problems with negative diagonal elements). If

$$\text{r.m.s. tr } |F| = \sqrt{\{\sum_{i=1}^{n} F(i,i)^2\}}$$

then,

$$\beta = \text{integer } \left| \log_2 \text{r.m.s. tr } |F| \right|$$

$$\alpha = 2^{\beta}$$

The problem encountered when using the simplified transition matrix definition and non-factored exponential series arise mainly from the system "blowing up" due to propagation of errors. Severe computational difficulties have been encountered in non-linear systems and although it is unclear as to which problem contributes most, it appears that the inclusion of the above refinements has been sufficient to eradicate these difficulties.

State dependency of measurement noise

The use of stochastic, dynamic systems in water quality modelling is useful in view of the almost inevitable presence of observation errors. Many of the observations are noise-corrupted by an amount which depends upon the magnitude of the true observation itself, i.e.

y = noise-corrupted observation

y_T = true value (i.e. what we should observe)

$\eta^* = y_T - y$

then often, η^* is a random function of y_T.

i.e. $\eta^* = \eta^*(y_T)$

The measurements of BOD and algae concentration in a river are cases where, for example the variance of the measurement error is proportional to y_T. Consider the measurement function of a continuous/discrete time system, where the measurements are assumed to be of this form, i.e.

$$\underline{y}(t_k) = h(\underline{x}(t_k), \underline{u}(t_k), t_k) + \underline{\eta}^*(t_k)$$

$$E(\underline{\eta}^*(t_k)\ \underline{\eta}^*(t_k)^T) = R_k^*$$

$$E(\underline{\eta}^*(t_k)) = 0$$

In the linear case, this becomes

$$\underline{y}(t_k) = H(t_k)\,\underline{x}(t_k) + \underline{\eta}^*(t_k)$$

Examining the filter innovation properties we find,

$$\underline{\nu}(t) = \underline{y}(t_k) - \underline{y}(t_k|t_{k-1})$$

$$= H(t_k)\,\underline{x}(t_k) + \eta^*(t_k) - H(t_k)\hat{\underline{x}}(t_k|t_{k-1})$$

$$= H(t_k)\tilde{\underline{x}}(t_k|t_{k-1}) + \underline{\eta}^*(t_k)$$

so that,

$$E(\underline{\nu}(t_k)) = E(H(t_k)\tilde{\underline{x}}(t_k|t_{k-1})) + E(\underline{\eta}^*(t_k))$$

$$= E(\underline{\eta}^*(t_k)) = 0$$

and

$$E(\underline{\nu}(t_k)\ \underline{\nu}(t_k)^T) = E((H(t_k)\tilde{\underline{x}}(t_k|t_{k-1}) + \underline{\eta}^*(t_k))(H(t_k)\tilde{\underline{x}}(t_k|t_{k-1}) + \underline{\eta}^*(t_k))^T)$$

Dropping the subscripts temporarily gives

$$E(\nu\nu^T) = E(H\tilde{x}\,\tilde{x}^T H^T + H\,\tilde{x}\,\eta^{*T} + \eta^*\,\tilde{x}^T H^T + \eta^*\,\eta^{*T})$$

We now define $\underline{\eta}^*$ as being a function of the state estimates, where

$$\underline{\eta}^*(t_k) \triangleq \omega(H(t_k)\hat{\underline{x}}(t_k|t_{k-1}))\underline{\eta}(t_k)$$

where

$$\omega(H(t_k)\hat{\underline{x}}(t_k|t_{k-1}))_{i,j} = \alpha_i \hat{y}_i(t_k|t_{k-1}),\ i = j$$

$$0,\ i \neq j$$

$$\underline{\eta}(t_k) \sim N(0,1)$$

This provides the link between the measurement noise and, in this case, an estimate of the observation. We can rewrite the innovations as,

$$E(\underline{\nu}(t_k)\underline{\nu}^T(t_k)) = E(H\,\tilde{\underline{x}}\,\tilde{\underline{x}}^T H^T) + H\,\tilde{\underline{x}}\,\underline{\eta}^T\,\omega^T(\hat{\underline{y}})$$
$$+ \omega(\hat{\underline{y}})\underline{\eta}\,\hat{\underline{x}}^T H^T + \omega(\hat{\underline{y}})\underline{\eta}\,\underline{\eta}^T\,\omega^T(\hat{\underline{y}}))$$

since

$$E(\tilde{\underline{x}}(t_k|t_{k-1})\ \underline{\eta}^T(t_k)) = 0.$$

Then $E(\underline{\nu}(t_k)\underline{\nu}^T(t_k)) = H(t_k)P(t_k|t_{k-1})H^T(t_k) + \omega(\hat{\underline{y}})\,R_k\,\omega^T(\hat{\underline{y}})$

where

$$R_k = E(\underline{\eta}(t_k)\ \underline{\eta}^T(t_k)) = I$$

and

$$E(\underline{\nu}(t_k)) = E(\underline{\eta}^*(t_k))$$

$$= E(\omega(\hat{\underline{y}}(t_k|t_{k-1}))\underline{\eta}(t_k))$$

$$= \omega(\hat{\underline{y}}(t_k|t_{k-1}))E(\underline{\eta}(t_k)) = 0$$

which demonstrates that the innovations are unbiased and that replacing R_k by $R_k^* = \omega(\underline{y})R_k\ \omega^T(\hat{\underline{y}})$ in the filter derivation enables us to amend the Kalman filter equations to

$$K(t_k) = P(t_k|t_{k-1})H^T(t_k)(H(t_k)P(t_k|t_{k11})H^T(t_k) + \delta(\hat{\underline{y}})R_k\,\omega^T(\hat{\underline{y}})]^{-1}$$

Thus, the linking of the measurement noise to the state estimates in the manner described does not destroy the optimality of the filter. By choosing $\omega(\underline{y})$ correctly, this can provide a weighting matrix which will sometimes, more accurately describe the behaviour of the real system.

B.2 User Notes for EKF program

PURPOSE: Filtering and prediction of non-linear, dynamic, stochastic, continuous/discrete state-space systems of the form;

$$\frac{d\underline{x}(t)}{dt} = f[\underline{x}(t), \underline{u}(t), t] + \xi(t)$$

$$\underline{y}(t_k) = H\,\underline{x}(t_k) + \underline{\eta}(t_k)$$

LANGUAGE: FORTRAN IV, single precision.

OPERATING SYSTEM: CDC

OUTPUT: On tape 6 (file OUP);

$\hat{\underline{x}}(t_k \mid t_k)$

$P(t_k|t_k)$

$\hat{\underline{x}}(t_{k+1}|t_k)$

$P(t_{k+1}|t_k)$

$\Phi(t_k; t_{k+1})$

$K(t_k)$

$\underline{y}(t_k) - \hat{\underline{y}}(t_k)$

INPUT: via tape 9 (file INP);

LINE 1 N M NDI NS

LINE 2 $\hat{\underline{x}}(t_k|t_{k-1})|_{t_k = t_o}$

LINE 3+ $P(t_k|t_{k-1})|_{t_k = t_o}$

LINE 4+ Q

LINE 5+ R

LINE 6+ H

LINE 7+ DELT

LINE 8 $\underline{u}(t_k)|_{t_k=t_o}$

LINE 9 $\underline{y}(t_k)|_{t_k=t_o}$

LINE 10 $\underline{u}(t_k)|_{t_k=t_o+\text{DELT}}$

LINE 11 $\underline{y}(t_k)|_{t_k=t_o+\text{DELT}}$

All variables are defined in the programme listing. Date is presented in free-format and all arrays are written column-wise for vectors, row-wise for matrices.

SUBROUTINES REQUIRED:

MODEL; contains a statement of the system equations.

Format:

```
SUBROUTINE MODEL (X, R, N, U, L, W, N4, PAR)
REAL X(N), U(L), W(N4), PAR(1Ø)
  .
  .
  .
return in the 1st N elements of W, the state differential equations
  .
  .
  .
RETURN
END
```

where:

X = state variable vector
R = real time value
N = number of states
U = deterministic input vector
L = number of deterministic inputs
W = vector of derivatives
N4 = 4 x N, length of W
PAR = utility vector

FDASH; contains a statement of the system Jacobian.

Format:

```
SUBROUTINE FDASH (F, X, U, N, L, R, PAR)
REAL F(N,N), X(N), U(L), PAR(1Ø)
  .
  .
  .
return in F the Jacobian equations
  .
  .
  .
RETURN
END
```

where:

F = Jacobian matrix

other notation as per MODEL

```
      PROGRAM EKF(INP,OUP,TAPE8=OUP,TAPE9=INP)
C
C     PROGRAMME EKF
C     ---------
C
C     Solves the filtering problem for any continuous/discrete
C     state-space system which is non-linear (or linear) in
C     the system equation and linear in the measurement
C     equation.
C
C     Dimension sizes:
C
C               n=number of states
C               m=number of observations
C               l=number of deterministic inputs
C
C     Variables:
C
C               TK=loop counter
C               N=n
C               M=m
C               NDI=l
C               NS=number of samples
C               DELT=constant sampling interval
C               A=beginning of integration step
C               B=end of integration step
C               STEP=initial value of integrator step length
C
C     Arrays:
C
C               XP=x(t+1|t)
C               XO=x(t|t)
C               PP=P(t+1|t)
C               PO=P(t|t)
C               Q=Q
C               R=R
C               H=measurement matrix
C               U=vector of deterministic inputs
C               Y=vector of observations
C               KG=Kalman gain matrix
C               ERRY=y(t+1)-H*x(t+1|t)
C               PHI=transition matrix
C               FSUM=Jacobian matrix
C               PAR=utility vector
C               W1,W2,W3,W4,W5=work arrays
C
C
C
      REAL XP(n),XO(n),PP(n,n),PO(n,n),Q(n,n),R(m,m)
      REAL H(m,n),U(l),Y(m),KG(n,m),ERRY(m)
```



```
      REAL PHI(n,n),FSUM(n,n),PAR(10)
      REAL W1(n,n),W2(n,n),W3(n,n),W4(n,n),W5(4n)
      INTEGER TK
C
C          input initialisation data
C
      READ(9,*)N,M,NDI,NS
C
C          initialise arrays
C
      DO10 I=1,N
      XP(I)=0.0
      XO(I)=0.0
10    ERRY(I)=0.0
      DO20 I=1,NDI
20    U(I)=0.0
      READ(9,*)(XP(I),I=1,N)
      DO30 I=1,N
30    READ(9,*)(PP(I,J),J=1,N)
      DO40 I=1,N
40    READ(9,*)(Q(I,J),J=1,N)
      DO50 I=1,M
50    READ(9,*)(R(I,J),J=1,M)

      DO60 I=1,M
60    READ(9,*)(H(I,J),J=1,N)
      READ(9,*)DELT
      A=0.0
      B=DELT
      STEP=DELT
C
C          commence main loop
C
      DO70 TK=1,NS
C
C          accept current input/ouput data
C
      READ(9,*)(U(I),I=1,NDI)
      READ(9,*)(Y(I),I=1,M)
C
C          evaluate gain matrix
C
      CALL GAIN(KG,H,PP,XP,R,N,M,W1,W2,W3)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
      WRITE(8,1000)A
      DO80 I=1,N
80    WRITE(8,1010)(KG(I,J),J=1,M)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
C
C          evaluate x(t|t)
C

C
      CALL COPRX(M,N,H,XP,ERRY,Y,KG,XO)
```

```
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
      WRITE(8,1020)A
      WRITE(8,1010)(ERRY(I),I=1,M)
      WRITE(8,1030)A,A
      WRITE(8,1010)(XO(I),I=1,N)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
C
C           evaluate P(t:t)
C
      CALL CORRP(KG,H,PP,PO,N,M,R,W1,W2)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
      WRITE(8,1040)A,A
      DO90 I=1,N
90    WRITE(8,1010)(PO(I,J),J=1,N)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
C
C           evaluate x(t+1:t)  and the Jacobian
C
      N4=4*N
      CALL DASCRU(A,B,STEP,N,XO,U,NDI,FSUM,PAR,N4,W5,W1,W2)
C
C     evaluate the transition matrix
C
      CALL FYE(FSUM,N,PHI,W1,W2,W3,W4)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
      WRITE(8,1050)A,B
      DO100 I=1,N
100   WRITE(8,1010)(PHI(I,J),J=1,N)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
      DO110 I=1,N
110   XP(I)=XO(I)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
      WRITE(8,1030)B,A
      WRITE(8,1010)(XP(I),I=1,N)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
C
C           evaluate P(t+1:t)
C
      CALL PREDP(PHI,PO,Q,PP,N,W1)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
      WRITE(8,1040)B,A
      DO110 I=1,N
110   WRITE(8,1010)(PP(I,J),J=1,N)
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
C
C           store data for residuals analysis
C
      CALL STATS(Y,ERRY,TK,M,NS)
      A=B
      B=B+DELT
```



```
70    CONTINUE
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
1000  FORMAT(1X,'K(',F7.3,')')
1010  FORMAT(15X,6(E11.4,1X))
1020  FORMAT(1X,'ERRY(',F7.3,')')
1030  FORMAT(1X,'X(',F7.3,',',F7.3,')')
1040  FORMAT(1X,'P(',F7.3,',',F7.3,')')
1050  FORMAT(1X,'PHI(',F7.3,';',F7.3,')')
C^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
      END
```

```
C
C
C
      SUBROUTINE GAIN(KG,H,PP,XP,R,N,M,W1,W2,W3)

C
C     Subroutine GAIN evaluates the Kalman gain matrix
C     according to:
C
C           K(t)=PP*H'*[H*PP*H'+R]**-1
C
C           H =H transpose
C
C
C     Variables:
C
C           SUM=work variable
C
C
C
      REAL KG(N,M),H(M,N),PP(N,N),XP(N),R(M,M)
      REAL W1(N,N),W2(N,N),W3(N,N)
C
C           evaluate PP*H'
C
      DO10 I=1,N
      DO10 J=1,M
      SUM=0.0
      DO20 K=1,N
20    SUM=SUM+PP(I,K)*H(J,K)
10    W1(I,J)=SUM
C
C           evaluate H*PP*H'
C
      DO30 I=1,M
      DO30 J=1,M
      SUM=0.0
      DO40 K=1,N
40    SUM=SUM+H(I,K)*W1(K,J)
30    W2(I,J)=SUM

C
C           evaluate H*PP*H'+R
C
      DO70 I=1,M
      DO70 J=1,M
70    W2(I,J)=W2(I,J)+R(I,J)
C
C           if Y is scalar, then avoid matrix inversion
C           in evaluation of [H*PP*H'+R]**-1
C
```



```
      IF(M.EQ.1)W3(1,1)=1.0/W2(1,1)
      IF(M.EQ.1)GOTO 80
      CALL INV(W2,W3,W4,N,M)
C
C           evaluate PP*H'*[H*PP*H'+R]**-1
C
80    DO90 I=1,N
      DO90 J=1,M
      SUM=0.0
      DO100 K=1,M
100   SUM=SUM+W1(I,K)*W3(K,J)
90    KG(I,J)=SUM
      RETURN
      END
```

```
C
C
C
      SUBROUTINE INV ( A,L,P,M,N)
C
C
C     Subroutine  INV inverts a  positive definite
C     symmetric  matrix by decomposing into lower
C     diagonal  form
C
C
C     Variables =
C
C          SUM = work variable
C          M=order of matrix to be inverted
C          N=DIMENSION of input s/output matrices
C
C
C     Arrays:
C
C          A=matrix to be inverted
C          L=lower diagonal,  returned inverse of A
C          P=work array
C
C
C
      REAL A(N,N),L(N,N),P(N,N)
C
C          initialise  arrays
C
      DO10 I=1,M
      DO10 J=1,M
      L(I,J)=0.0
10    P(I,J)=0.0
C
C          check for  +ve definiteness
C
      IF(A(1,1).LE.0.0)GOTO 999
C
C          transform  A into lower diagonal form L
C
      L(1,1)=SQRT(A(1,1))
      DO20 I=2,M
20    L(I,1)=A(1,I)/L(1,1)
      DO30 I=2,M
      SUM=0.0
      DO40 K=1,I-1
40    SUM=SUM+L(I,K)*L(I,K)
C
C          check for  +ve definiteness
C
C
```

```
      CHECK=A(I,I)-SUM
      IF(CHECK.LE.0.0)GOTO 999
      L(I,I)=SQRT(CHECK)
      IF(I.EQ.M)GOTO 30
      DO50 J=I+1,M
      SUM=0.0
      DO60 K=1,I-1
60    SUM=SUM+L(J,K)*L(I,K)
50    L(J,I)=(A(I,J)-SUM)/L(I,I)
30    CONTINUE
C
C          evaluate L**-1
C
      DO70 I=1,M
70    P(I,I)=1.0/L(I,I)
      DO80 I=2,M
      DO80 J=1,I-1
      SUM=0.0
      DO90 K=J,I-1
90    SUM=SUM+L(I,K)*P(K,J)
80    P(I,J)=-SUM/L(I,I)
      DO100 I=1,M-1
      DO100 J=I+1,M
100   P(I,J)=0.0
C
C          evaluate A**-1=L**-1t*L**-1
C
      DO120 I=1,M
      DO120 J=1,M
      SUM=0.0
      DO130 K=I,M
130   SUM=SUM+P(K,I)*P(K,J)
120   L(I,J)=SUM
C
C          error sequence
C
      RETURN
999   WRITE(6,1000)
1000  FORMAT(1X,'matrix H*P*H'+R is non +ve definite')
      STOP
      END
```

```
C
C
C
      SUBROUTINE CORRX(M,N,H,XP,ERRY,Y,KG,XO)
C
C
C     Subroutine CORRX evaluates corrected state
C     estimate according to:
C
C           x(t|t)=x(t|t-1)+KG*[Y(t)-H*x(t|t-1)]
C
C     Variables:
C
C           SUM=work variable
C
C
      REAL H(M,N),XP(N),ERR(M),Y(M),KG(N,M),XO(N)
C
C     evaluate err(y)=Y(t)-H*x(t|t-1)
C
      DO10 I=1,M
      SUM=0.0
      DO20 J=1,N
20    SUM=SUM+H(I,J)*XP(J)
10    ERRY(I)=Y(I)-SUM
C
C     evaluate x(t|t)
C
      DO30 I=1,N
      SUM=0.0
      DO40 J=1,M
40    SUM=SUM+KG(I,J)*ERRY(J)
30    XO(I)=XP(I)+SUM
      RETURN
      END
```



```
C
C
C
      SUBROUTINE CORRP(KG,H,PP,PO,N,M,R,W1,W2)
C
C
C     Subroutine CORRP evaluates the predicted state
C     estimate error covariance matrix according to:
C
C           P(t+1|t)=[I-K(t)*H]*P(t|t)*[I-K(t)*H]'
C                   +K(t)*R*K'(t)
C
C
C     Variables:
C
C           SUM=work variable
C
      REAL R(M,M),KG(N,M),H(M,N),PP(N,N),PO(N,N)
      REAL W1(N,N),W2(N,N)
C
C     evaluate I-K(t)*H
C
      DO10 I=1,N
      DO10 J=1,N
      SUM=0.0
      DO20 K=1,M
20    SUM=SUM+KG(I,K)*H(K,J)
10    W1(I,J)=SUM
      DO30 I=1,N
      DO30 J=1,N
      W2(I,J)=0.0
30    W2(I,I)=1.0
      DO40 I=1,N
      DO40 J=1,N
40    W2(I,J)=W2(I,J)-W1(I,J)
C
C     evaluate [.]*P(t|t)*[.]'
C
      DO50 I=1,N
      DO50 J=1,N
      SUM=0.0
      DO60 K=1,N
60    SUM=SUM+W2(I,K)*PP(K,J)
50    W1(I,J)=SUM
      DO70 I=1,N
      DO70 J=1,N
      SUM=0.0
      DO80 K=1,N
80    SUM=SUM+W1(I,K)*W2(J,K)
```

```
 70   PO(I,J)=SUM
C
C        evaluate K*R*K'
C
      DO90 I=1,N
      DO90 J=1,M
      SUM=0.0
      DO100 K=1,M
100   SUM=SUM+KG(I,K)*R(K,J)
 90   W1(I,J)=SUM
      DO110 I=1,N
      DO110 J=1,N
      SUM=0.0
      DO120 K=1,M
120   SUM=SUM+W1(I,K)*KG(J,K)
110   PO(I,J)=PO(I,J)+SUM
      RETURN
      END
```



```
C
C
C
      SUBROUTINE FYE(FD,N,PHI,W1,W2,W3,W4)
C
C
C     Subroutine FYE evaluates the matrix exponential using
C     a truncated series expansion.
C
C
C     Variables:
C
C          TRACE=root mean square trace of FD
C          FLAG=flag to detect factoring
C          COUNT=log(base2) of TRACE
C          DIV=2**(integer protion of COUNT)
C          NT=maximum number of terms in series
C          C=counter for factorial
C          X=recursive factorial element
C          XM=largest element of current term
C          ERR=piecewise absolute elements of current term
C          SUME=work variable
C
C
C     Arrays:
C
C          FD=matrix to be exponentiated
C          PHI=resultant exponential
C
C
C
      REAL FD(N,N),PHI(N,N)
      REAL W1(N,N),W2(N,N),W3(N,N),W4(N,N)
C
C     evaluate root mean square trace
C
      TRACE=0.0
      DO10 I=1,N
 10   TRACE=TRACE+FD(I,I)*FD(I,I)
      TRACE=SQRT(TRACE)
      FLAG=0.0
      IF(TRACE.LE.2.0)GOTO 20
C
C     if TRACE > 2 then factor FD
C
      COUNT=ALOG(TRACE)/0.69314718
      COUNT=INT(COUNT)
      DIV=2.0**COUNT
      DO30 I=1,N
      DO30 J=1,N
```

```
30      FD(I,J)=FD(I,J)/DIV
20      CONTINUE
C
C            initialise arrays
C
        DO40 I=1,N
        DO50 J=1,N
        PHI(I,J)=0.0
        W1(I,J)=0.0
        W2(I,J)=0.0
50      W3(I,J)=0.0
        W1(I,I)=1.0
40      PHI(I,I)=1.0
        C=0.0
C
C            compute factorial element
C
55      C=C+1
        X=1.0/C
        DO60 I=1,N
        DO60 J=1,N
        W2(I,J)=0.0
60      W3(I,J)=FD(I,J)*X
        XN=0.0
        DO70 I=1,N
        DO80 K=1,N
        DO80 J=1,N
80      W2(I,K)=W2(I,K)+W1(I,J)*W3(J,K)
70      CONTINUE
C
C            add term to result
C
        DO90 I=1,N
        DO90 J=1,N
        PHI(I,J)=PHI(I,J)+W2(I,J)
C
C            check magnitude of current term
C
        ERR=ABS(W2(I,J))
        W1(I,J)=W2(I,J)
90      IF(ERR.GT.XN)XN=ERR
C
C            check for end of series
C
        IF(C.GT.NT)GOTO 100
C
C            check current term against desired accuracy
C
        IF(XN.GT.1.0E-6)GOTO 55
100     CONTINUE
```



```
        IF (TRACE.LE.2.0)GOTO 110
C
C            if FD was factored.  invert factorisation by
C            repeated squaring
C
120     DO1 30 I=1,N
        DO1 30 J=1,N
        SUME=0.0
        DO1 40 K=1,N
140     SUME=SUME+PHI(I,K)*PHI(K,J)
130     W4(I,J)=SUME
        DO1 50 I=1,N
        DO1 50 J=1,N
150     PHI(I,J)=W4(I,J)
        IF(COUNT.LT.1.5)GOTO 110
        COUNT=COUNT-1.0
        GOTO 120
110     CONTINUE
        RETURN
        END
```

```
C
C
C
      SUBROUTINE STATS(Y,ERRY,TK,M,NS)
C
C
C           Subroutine STATS evaluates basic statistics
C           of the residuals sequences.
C
C
C           Variables:
C
C                 EN=REAL number of samples
C
C           Arrays:
C
C                 S=array of sums
C                 S2=array of sum squares
C                 MEANR=vector of residual means
C                 MEANY=  "      "  observation means
C                 VARR=vector of residual variances
C                 VARY=  "      "  observation variances
C                 R2=vector of coefficients of variation
C
C
C
C
      REAL S(2,m),S2(2,m),MEANR(m),MEANY(m),VARR(m),VARY(m),R2(m)
      REAL Y(M),ERRY(M)
      INTEGER TK
      EN=FLOAT(NS)
      DO10 I=1,M
      IF(TK.NE.1)GOTO 20
C
C           initialise arrays
C
      DO30 J=1,M
      DO30 K=1,2
      S(K,J)=0.0
30    S2(K,J)=0.0
20    CONTINUE
C
C           evaluate sums and sum squares
C
      S(1,I)=S(1,I)-ERRY(I)
      S(2,I)=S(2,I)+Y(I)
      S2(1,I)=S2(1,I)+ERRY(I)*ERRY(I)
      S2(2,I)=S2(2,I)+Y(I)*Y(I)
      IF(TK.LT.NS)GOTO 10
C
C           evaluate means and variances
C
C
```



```
      MEANR(I)=S(1,I)/EN
      MEANY(I)=S(2,I)/EN
      VARR(I)=(S2(1,I)-EN*MEANR(I)*MEANR(I))/(EN-1.0)
      VARY(I)=(S2(2,I)-EN*MEANY(I)*MEANY(I))/(EN-1.0)
      R2(I)=(VARY(I)-VARR(I))/VARY(I)
10    CONTINUE
      IF(TK.LT.NS)RETRUN
C
C           write results
C
      WRITE(8,1000)NS
      WRITE(8,1010)(MEANR(I),I=1,M)
      WRITE(8,1020)(VARR(I),I=1,M)
      WRITE(8,1030)(MEANY(I),I=1,M)
      WRITE(8,1040)(VARY(I),I=1,M)
      WRITE(8,1050)(R2(I),I=1,M)
      RETURN
1000  FORMAT(1X,'number of samples ',I5/1X,'mean residuals'/)
1010  FORMAT(1X,6(E10.3,2X))
1020  FORMAT(1X,'variance residuals'/6(2X,E10.3))
1030  FORMAT(1X,'mean observations'/6(2X,E10.3))
1040  FORMAT(1X,'variance observations'/6(2X,E10.3))
1050  FORMAT(1X,'coefficients of variation'/6(2X,E10.3))
      END
```

```
C
C
C
      SUBROUTINE DASCRU(A,B,H,N,XO,U,NDI,FSUM,PAR,N4,WK,W1,W2)
C
C
C     Subroutine DASCRU is a modified library routine
C     which integrates a set of first-order ordinary
C     differential equations by the Runge-Kutta method
C     and evaluates the integral Jacobian matrix.
C
C
C     Variables:
C
C          E5=1/2 desired accuracy of solution
C          H=initial guess at step length
C          HMIN=minimum step length
C
C     Arrays:
C
C          XO=state vector x
C          FSUM=Jacobian matrix
C
C
C     Requires two subroutines, MODEL and FDASH.
C     MODEL defines the set of differential equations,
C     FDASH defines the Jacobian matrix.
C
C
C     DASCRU halves H repeatedly until either the desired
C     accuracy is achieved or H=HMIN. If H is set initially
C     too small, DASCRU doubles H.
C
C
C
      REAL WK(N4),XO(N),U(NDI),FSUM(N,N),W1(N,N)
      REAL PAR(10),W2(N,N)
      INTEGER SW
      LOGICAL BE,BH,BR,BX,I,J
      DATA ZERO,P5,OP5,THREE,FOUR,E5/0.0,0.5,1.5,3.0,4.0,0.5E-4/
      DO202 I=1,N
      DO202 J=1,N
202   FSUM(I,J)=0.0
      X=A
      CALL FDASH(W1,XO,U,N,NDI,X,PAR)
      IF(A-B)4,100,4
4     IB1=N+N
      IB2=IB1+N
      HMIN=1.0E-2*ABS(H)
      BH=.TRUE.
```



```
      BR=.TRUE.
      BX=.TRUE.
      H=SIGN(ABS(H),B-A)
      X=A
5     XS=X
      DO 10 J=1,N
      IJKO=N+J
      WK(IJKO)=XO(J)
10    CONTINUE
15    HS=H
      Q=X+H-B
      BE=.TRUE.
      IF(.NOT.((H.GT.ZERO.AND.Q.GE.ZERO).OR.(H.LT.ZERO.AND.Q.LE.
     1ZERO)))GO TO 20
      H=B-X
      BR=.FALSE.
20    H3=H/THREE
      DO 90 SW=1,5
      DO 120 I=1,N
120   WK(I)=0.0
      CALL MODEL(XO,X,N,U,NDI,WK,N4,PAR)
      DO 70 I=1,N
      Q=H3*WK(I)
      IJKO=N+I
      IJK1=IB1+I
      IJK2=IB2+I
      GO TO (25,30,35,40,45),SW
25    R=Q
      WK(IJK1)=Q
      GO TO 50
30    R=P5*(Q+WK(IJK1))
      GO TO 50
35    R=THREE*Q
      WK(IJK2)=R
      R=0.375*(R+WK(IJK1))
      GO TO 50
40    R=WK(IJK1)+FOUR*Q
      WK(IJK1)=R
      R=OP5*(R-WK(IJK2))
      GO TO 50
45    R=P5*(Q+WK(IJK1))
      Q=ABS(R+R-OP5*(Q+WK(IJK2)))
50    XO(I)=WK(IJKO)+R
      IF(SW.NE.5)GO TO 70
      E=ABS(XO(I))
      R=E5
      IF(E.GE.1.0E-3)R=E*E5
      IF(Q.LT.R.OR.(.NOT.BX))GO TO 65
      BR=.TRUE.
      BH=.FALSE.
```

```
      H=P5*H
      IF(ABS(H).GE.HMIN)GO TO 55
      H=SIGN(1.0,H)*HMIN
      BX=.FALSE.
55    DO 60 J=1,N
      IJKO=N+J
      XO(J)=WK(IJKO)
60    CONTINUE
      X=X5
      GO TO 15
65    IF(Q.GE.0.03*.25*R)BE=.FALSE.
70    CONTINUE
      GO TO (75,90,80,85,90),SW
75    X=X+H3
      GO TO 90
80    X=X+P5*H3
      GO TO 90
85    X=X+P5*H
90    CONTINUE
      CALL FDASH(W2,XO,U,N,NDI,X,PAR)
      DO 201 I=1,N
      DO 201 J=1,N
201   FSUM(I,J)=FSUM(I,J)+(W1(I,J)+W2(I,J))*H/2.0
      DO 204 I=1,N
      DO 204 J=1,N
204   W1(I,J)=W2(I,J)
      IF(.NOT.(BE.AND.BH.AND.BR))GO TO 95
      H=H+H
      BX=.TRUE.
95    BH=.TRUE.
      IF(BR)GO TO 5
      H=HS
      RETURN
100   STOP
      END
```



```
C
C
C
      SUBROUTINE PREDP(PHI,PO,Q,PP,N,W1)

C     Subroutine PREDP evaluates the predicted state
C     estimate error covariance matrix according to:
C
C           P(t+1|t)=PHI*P(t|t)*PHI'+Q
C
C     Variables:
C
C           SUM=work variable
C
C
      REAL PHI(N,N),PO(N,N),PP(N,N),Q(N,N),W1(N,N)
C
C     evaluate P(t|t)*PHI'
C
      DO10 I=1,N
      DO10 J=1,N
      SUM=0.0
      DO20 K=1,N
20    SUM=SUM+PO(I,K)*PHI(J,K)
10    W1(I,J)=SUM
C
C           evaluate PHI*P(t|t)*PHI' + Q
C
      DO30 I=1,N
      DO30 J=1,N
      SUM=0.0
      DO40 K=1,N
40    SUM=SUM+PHI(I,K)*W1(K,J)
30    PO(I,J)=SUM+Q(I,J)
      RETURN
      END
```

```
C
CC
C
      SUBROUTINE MODEL(X,R,N,U,L,W,N4,PAR)
C
CCCC
CC
C
      Example MODEL routine

      REAL X(N),U(L),W(N4),PAR(10)
      DO10 I=1,N
10    W(I)=0.0
      W(1)=0.33*X(7)*(8.65-X(1))
     1    -2.142E-5*X(5)*X(3)*X(2)
     2    -5.06E-7*X(6)*X(2)*X(2)/X(3)
     3    -0.21*X(1)+0.14
      W(2)=1.0E-3*X(5)*X(3)*X(2)
     1    -1.0E-4*X(6)*X(2)*X(2)/X(3)
     2    -0.21*X(2)+700.0
      W(3)=-1.0E-4*X(3)*X(4)*X(2)
     1    -0.21*X(3)+0.14*U(1)
      RETURN
      END
```



```
C
CC
C
      SUBROUTINE FDASH(F,X,U,N,L,R,PAR)
C
CC
CCCC
C
      example FDASH routine

      REAL F(N,N),X(N),U(L),PAR(10)
      DO10 I=1,N
      DO10 J=1,N
10    F(I,J)=0.0
      F(1,1)=-0.33*X(7)-0.21
      F(1,2)=-2.142E-5*X(5)*X(3)
     1       -2.0*5.06E-7*X(6)*X(2)/X(3)
      F(1,3)=-2.142E-5*X(5)*X(2)
     1       +5.06E-7*X(6)*X(2)*X(2)/(X(3)*X(3))
      F(1,5)=-2.142E-5*X(3)*X(2)
      F(1,6)=-5.06E-7*X(2)*X(2)/X(3)
      F(1,7)=0.33*(8.65-X(1))
      F(2,2)=1.0E-3*X(5)*X(3)
     1       -1.0E-4*2.0*X(6)*X(2)/X(3)-0.21
      F(2,3)=1.0E-3*X(5)*X(2)
     1       +1.0E-4*X(6)*X(2)*X(2)/(X(3)*X(3))
      F(2,5)=1.0E-3*X(2)*X(3)
      F(2,6)=-1.0E-4*X(2)*X(2)/X(3)
      F(3,2)=-1.0E-4*X(3)*X(4)
      F(3,3)=-1.0E-4*X(4)*X(2)-0.21
      F(3,4)=-1.0E-4*X(3)*X(2)
      RETURN
      END
```