DATA-BASED MECHANISTIC MODELLING (DBM) OF NONLINEAR ENVIRONMENTAL SYSTEMS

Christopher P. Fawcett, B.Sc. (Hons.)

A thesis submitted to the University of Lancaster for the degree of Doctor of Philosophy

August 1999

Environmental Science Department
Institute of Environmental and Natural Sciences
Lancaster University, United Kingdom

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

DECLARATION

The experimental work described in this thesis was original work by the author undertaken at the Department of Environmental Science, Lancaster University. The material has not been submitted, in whole or in part for any degree, to this or any other institution.

Work from this thesis also appears in articles listed in appendix 1.

DATA-BASED MECHANISTIC MODELLING OF NONLINEAR ENVIRONMENTAL SYSTEMS.

Submitted by

Christopher P. Fawcett, B.Sc. (Hons.).

A thesis submitted to the University of Lancaster

for the degree of Doctor of Philosophy

August 1999.

ABSTRACT

The main focus of the research studies presented in this thesis centre on the application and development of data-based mechanistic (DBM) transfer function models for nonlinear environmental systems. The data-based mechanistic modelling approach exploits the available time series data, in statistical terms, to expose the model structure, generating dynamic stochastic models that are parsimonious in nature and can be interpreted in a physical manner.

For nonlinear systems, the DBM approach centres on the use of transfer function models whose parameters are free to vary over time. The presence of such time variable parameters may reflect either nonstationarity or nonlinearity; the latter arising if the variations are also shown to be state dependent. Statistical time series methods for estimating time varying parameters (TVP) and the use of these methods in state dependent parameter modelling (SDPM) are discussed and applied to the modelling of nonlinear ecological population dynamics and hydrological processes. Further, DBM

modelling techniques are applied to an oceanic ecosystem simulation model in order to investigate model uncertainty and over-parameterisation, set within the context of data assimilation.

In each application, the DBM methodology is shown to successfully identify the system nonlinearities, so providing additional physical insight and ensuring a good explanation of the data with the minimum of model parameters (parsimony). Further, the DBM approach provides an approach to both the evaluation and reduction in the complexity of highly parameterised simulation models.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Professor Peter Young and Dr Arun Chotai for their supervision and constant encouragement throughout the course of these research studies. Dr John Rowan is also thanked for his guidance and enthusiasm, and the assistance from Dr Mike Fasham is also appreciated.

Present and past members of the Systems and Control research group are also mentioned for providing such welcome humorous relief from the computer terminals. Paul deserves particular thanks for his helpful advice and suggestions. Mark, Glenda, David, Simon and other colleagues from IENS are also thanked.

I am also grateful for the continued words of encouragement from friends and family over the last four years. But most importantly, I would like to thank Laura for her continued, support, belief and patience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Mathematical modelling spans a wide range of philosophies and methodological approaches. Models of environmental systems are predominantly deterministic in nature which, in line with current paradigm, favours models that provide a physically interpretable structure. An alternative systems approach, which has its origins in the field of engineering and statistics, provides models that are developed directly from time series data. Environmental modellers are beginning to embrace this *data-based* modelling approach, realising the value of the enhanced statistical model estimation tools and the parsimonious nature of data-based models.

Deterministic or mechanistic mathematical models are generally formulated to have a structure that closely resembles the physical, chemical or biological reality of the system as *perceived* by the modeller using classical dynamic conservation equations (mass, energy and momentum) and are, therefore, heavily reductionist in nature. As a consequence, these models are limited by the modeller's experience and overall understanding of the system and often become extremely complex and nonlinear as they attempt to describe the physical system in some detail, reflecting the complexity

that the modeller perceives in nature. Furthermore, as the parameters of deterministic models often have a perceived physical interpretation, the feasibility, cost and time in obtaining sufficiently informative data for their estimation, result in the models becoming extremely difficult to validate.

Conversely, purely data-based (*black box* or *time series*) models are generated without any prior assumptions regarding the systems internal behaviour. The model's structure is identified and parameters estimated objectively from observational time series data using appropriate statistically-based methods. As such, data-based models are simple in structure and parsimonious; characterised by only the number of parameters that can be justified by the data. While such statistical models have the advantage of being data-based and stochastic, they are dependent upon the availability and quality of data and are, consequently, often unsuitable for simulation purposes. Furthermore, because they do not attempt to relate physically to the system they surmise, many scientists consider the overall value of these types of model as low and reject there utility on physical grounds.

The data-based mechanistic (DBM) modelling approach is a *hybrid* of these two modelling extremes. Rejecting the deterministic-reductionist philosophy, DBM models retain the statistical advantages and parsimony of the data-based modelling approach but address the desire for a physical interpretation of the model. In the majority of cases, the DBM methodology has been applied to linear systems, or where nonlinear systems have been approximated by linear models. In the main, this has been due to a restriction of computing capabilities, limiting the development of more complex nonlinear identification and estimation procedures. However, in this thesis,

the DBM philosophy is invoked with a focus on the modelling of nonlinear systems using techniques that have been developed more recently (Young, 1984; 1988; 1993) and allow for the effective identification and estimation of nonlinear models.

This chapter introduces the DBM philosophy, with a particular emphasis placed on nonlinear systems, and presents the novel contributions of this thesis, where the DBM methodology and techniques are applied to model several different environmental systems.

1.1 DATA-BASED MECHANISTIC MODELLING

The data-based mechanistic (DBM) modelling methodology is designed to be as objective as possible, exploiting the available time series in statistical terms to expose the model's structure, rather than being based on prior assumptions, in order to generate dynamic stochastic models that are parsimonious in nature. DBM modelling is based around the simple input-output transfer function (TF) model, originally derived in the control engineering literature and later exploited in discrete-time terms in the statistical literature by Box and Jenkins (1970). Here, the TF model structure is identified and its parameters estimated directly from the data using statistical methods. This allows for any uncertainty associated with both the model and the noisy data; and it ensures that the DBM model reflects only the dominant modes of the system behaviour.

For nonlinear systems, the DBM modelling approach is centred around the utilisation of transfer function (TF) models whose parameters are free to vary over time and are

estimated from the time series by a powerful recursive Fixed Interval Smoothing (FIS) algorithm (see e.g. Young, 1984; 1993). In this setting, any parameters that exhibit significant temporal variation over the observation interval will reflect either non-stationarity in the system (i.e. its characteristics change over time) or a nonlinearity in the system dynamics. The modelling proceeds to determine whether any identified time varying parameter can be related to any states of the system, so allowing for its direct parameterisation into the model. If the nonlinearity has been effectively identified and parameterised by the 'State Dependent Parameter Model' (SDPM), then the resulting nonlinear TF model will usually have time invariant parameters. A final prerequisite of the DBM philosophy requires that, in addition to explaining the data well, the final identified DBM model is credible in physical terms; in other words, it is only accepted as a truly useful model of the system if its structure and parameters can be interpreted in a physical manner.

As the DBM methodology is truly generic in its approach, it has been utilised over a wide range of disciplines over the last 10 years, including the fields of engineering, economics, biology, and environmental science. As an example of the range of DBM applications: Lees *et al.* (1994) developed an online DBM flood forecasting model, which utilises recursive time varying parameter estimation, for the River Nith in Dumfriesshire; while at the other extreme, Jarvis *et al.*, (1998) employed the DBM modelling approach to investigate the dynamic response of plant stomatal conductance to the reduction in atmospheric humidity. However, most of these applications consider linear or near-linear systems. In this thesis, the same approach is directed towards the identification and estimation of nonlinear, stochastic dynamic systems.

1.1.1 Transfer function (TF) models

Over the past 30 years, the extensive development of digital computing has increasingly encouraged the development of systems models and theories expressed in discrete-time terms, as opposed to continuous-time theories which tended to be associated with analogue computing. Discrete-time models are based upon sampled data and are often presented in terms of either the forward shift, z^i, or backward shift z^{-i} operators which are defined as,

$$z^i x(k) = x(k+i) \tag{1.1}$$

$$z^{-i} x(k) = x(k-i) \tag{1.2}$$

where $x(k)$ is a signal at the kth sampling instant. The transfer function model, which derives from the use of such operators, is simple but effective in its approach, relating an input signal to an output signal which, in discrete-time, is normally assumed to be sampled regularly at a sampling interval of Δt time units. The discrete-time transfer function model of a single-input, single-output (SISO) system takes the form,

$$y(k) = \frac{B(z^{-1})}{A(z^{-1})} u(k-\Delta) \tag{1.3}$$

where $u(k)$ and $y(k)$ are the discretely sampled input and output signals, respectively, at the kth sampling instant (e.g. values sampled after $k\Delta t$ time units). Here, $A(z^{-1})$ and $B(z^{-1})$ are the polynomials in the z^{-1} operator, defined as follows.

$$A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_n z^{-n} \tag{1.4}$$

$$B(z^{-1}) = b_0 + b_1 z^{-1} + \cdots + b_m z^{-m} \tag{1.5}$$

The integers n and m are the number of parameters in polynomials (1.4) and (1.5) respectively. No prior assumptions are made about the nature of the TF, which may be marginally stable, unstable or possess non-minimum phase characteristics. Any pure time delay in the system, which often characterises environmental systems, appears in equation (1.3) as the delay of Δ time intervals associated with the input $u(k-\Delta)$.

In this linear model (1.3), two physically meaningful properties can be derived from the transfer function model if its eigenvalues are real and the model is stable: the *steady state gain* and *time constant*. The steady state gain, G, defines the relationship between the equilibrium output value when a constant input is applied to the system. G is calculated by setting the z^{-1} operator in equation (1.3) to unity such that, in the case of a first order TF,

$$G = \frac{b_0}{1 + a_1} \tag{1.6}$$

As a result, the steady state gain can be used to determine whether there has been any physical gains or losses to the system. The time constant or residence time T_c is the time required for the system output to decay to e^{-1} (~37%) of its maximum value in response to a unit impulse and is defined as,

$$T_c = \frac{-1}{\log_e(-a_1)} \tag{1.7}$$

Where the polynomials in equation (1.3) have multiple-order (e.g. third order), the TF model can sometimes be unambiguously decomposed into lower order (e.g. first or second order) TF elements connected in serial, parallel or feedback form. For example, a second order transfer function ($m=1$ and $n=2$ in (1.4) and (1.5) respectively) can be represented by two first-order TF models connected either in parallel or in feedback. As such, the decomposition of multi-order TF models can sometimes reveal useful information about the system's internal behaviour that would otherwise remain unknown.

In its present form, the constant parameter TF model defined by equation (1.3) is suitable for modelling linear systems. As discussed in detail in Chapter 2, equation (1.3) can be conveniently redefined into an alternative time varying parameter (TVP) or state dependent parameter TF form, suitable for modelling systems that are either non-stationary or nonlinear.

The approach to identifying systems using time variable parameter estimation used in this thesis has its origins in the early 1970's (e.g. Beck and Young, 1976), where the Extended Kalman Filter (EKF) was used to infer the structural nature of models for biochemical oxygen demand and dissolved oxygen in rivers. Since then, the use of the EKF in this role has been explored in detail by Beck in a number of papers and chapters in books e.g. Beck, (1985); Beck, (1987); Kleissen *et al.*, (1990); Chen and Beck, (1993); and Eigbe *et al.*, (1998). This approach has similarities, in a

philosophical sense, with the alternative approach used in this thesis, although the methodology apart from the exploitation of recursive estimation, is quite different.

Equation (1.8) below presents the general form of the time varying parameter TF model. The nonlinear relationship existing between the system input $u(k)$ and output $y(k)$ is reflected by the time varying nature of the parameters.

$$y(k) = \frac{B(k, z^{-1})}{A(k, z^{-1})} u(k) = \frac{b_0(k) + b_1(k) z^{-1} + \cdots + b_m(k) z^{-m}}{1 + a_1(k) z^{-1} + \cdots + a_n(k) z^{-n}} u(k) \tag{1.8}$$

In the situation where the model parameters vary rapidly (i.e. at a rate consecutive with the variations in the states of the system itself), the behaviour of the time varying parameters of TF (1.8) may be characterised by potentially nonlinear functions of related variables that are present within a suitably defined nonminimum state space (NMSS) vector $\boldsymbol{\chi}(k)$. The resulting general form of this state dependent parameter TF model is defined as,

$$\begin{aligned} y(k) &= \frac{B(\{\boldsymbol{\chi}(k)\}, z^{-1})}{A(\{\boldsymbol{\chi}(k)\}, z^{-1})} u(k) \\ &= \frac{b_0\{\boldsymbol{\chi}(k)\} + b_1\{\boldsymbol{\chi}(k)\} z^{-1} + \cdots + b_m\{\boldsymbol{\chi}(k)\} z^{-m}}{1 + a_1\{\boldsymbol{\chi}(k)\} z^{-1} + \cdots + a_n\{\boldsymbol{\chi}(k)\} z^{-n}} u(k) \end{aligned} \tag{1.9}$$

where the NMSS vector $\boldsymbol{\chi}(k)$ is defined as,

$$\boldsymbol{\chi}(k) = \left[y(k-1), \cdots, y(k-n) \boldsymbol{u}(k)^{\mathrm{T}}, \cdots, \boldsymbol{u}(k-m)^{\mathrm{T}} \boldsymbol{U}(k)^{\mathrm{T}}, \cdots, \boldsymbol{U}(k-q)^{\mathrm{T}} \right]^{\mathrm{T}} \tag{1.10}$$

and the terms $b_0\{\boldsymbol{\chi}(k)\}$, $b_1\{\boldsymbol{\chi}(k)\}$, $b_m\{\boldsymbol{\chi}(k)\}$, $a_1\{\boldsymbol{\chi}(k)\}$, $a_n\{\boldsymbol{\chi}(k)\}$ indicate that the parameters are nonlinear functions of this NMSS vector. Here, $\boldsymbol{\chi}(k)$ constitutes a

vector of past values of the system output $y(k)$; present and past values of the deterministic input variables $\boldsymbol{u}(k)$ that are considered to be the *principle* inputs to the system; and vector the $\boldsymbol{U}(k)$, which contains present and past values of additional input variables that *may* affect the system nonlinearly. The thesis will argue that this kind of 'State Dependent Parameter Model' (SDPM) can provide a generic model for the identification and estimation of a wide class of nonlinear systems within the environmental and ecological sciences.

1.2 GENERAL SCOPE OF THIS THESIS

A summary of the overall aims of this thesis is given below.

- This thesis emphasises the use of DBM methodology as a tool for modelling nonlinear environmental systems. The techniques of time varying parameter (TVP) estimation and state dependent parameter modelling (SDPM) which are fundamental to objective nonlinear model development, are applied throughout this thesis.

- The DBM methodology is applied to environmental applications, specifically modelling of nonlinear ecological and hydrological systems.

- The DBM modelling philosophy generally rejects the use of large deterministic models for systems where there is an abundant of quality observational data. However, the use of much larger simulation models, which reflect the current scientific paradigm, is also a crucial requirement for systems where little data

are available. But we believe that such simulation models must acknowledge the uncertainty of the environmental system under consideration. Consequentially, the third aim of this thesis is to show how stochastically defined simulation models and data-based statistical methods can be used, in combination, for the analysis of environmental systems that are poorly defined. In this manner, it is possible to identify the strengths and limitations of the simulation model, investigate alternative modelling approaches, and provide more information about the system under consideration.

1.3 STRUCTURE OF THIS THESIS

The present chapter has briefly introduced the DBM methodology which will be used throughout this thesis in different modelling applications. Chapter 2 discusses the main statistical methods for the identification and estimation of linear and nonlinear discrete-time models that are used in these applications. In particular, the following two algorithms are described in detail: the Simplified Refined Instrumental Variable (SRIV) algorithm (Young, 1984); and the recursive Fixed Interval Smoothing (FIS) algorithm for time varying parameter estimation (see e.g. Young, 1984). Furthermore, deterministic and stochastic methods for optimising the parameters of the final nonlinear model are introduced, including the Maximum Likelihood method (see e.g. Harvey, 1989).

Chapter 3 describes how the DBM modelling methodology can be used to effectively model the nonlinear behaviour of the population dynamics in the data series of the well known Nicholson sheep blowfly experiments (Nicholson, 1950; 1954; 1957).

In Chapter 4, the Lancaster DBM nonlinear rainfall-flow model (Young, 1993; Young and Beven, 1994) which incorporates an effective rainfall term as a physical indicator of catchment antecedent soil moisture conditions, is validated for the first time using groundwater table and soil moisture measurements. A comparison between the nonlinear components of the Lancaster DBM and IHACRES rainfall-flow models (Jakeman *et al.*, 1990a) is made, with reference to field data collected by ETH, Zurich. These data justify the form of the nonlinear component in the Lancaster model, which is identified using DBM techniques and provides an alternative form of the model where such additional data are available.

Chapter 5 investigates how the DBM methodology has been be applied to the small upland catchment of Wyresdale Park, Lancashire, in order to extend contemporary time series of suspended sediment load. A discrete-time TF model for the rainfall-sediment relationship is identified and estimated from 2 years of field data. This model is used, together with historical rainfall series, to simulate suspended sediment load series over the past 100 years in order to generate a synthetic sediment flood sequence for the catchment reservoir.

Chapter 6 presents a critical evaluation of a deterministic oceanic ecosystem simulation model with the overall aim of incorporating it into a data assimilation (Robinson *et al.*, 1999) framework. By considering the deterministic simulation equations of the system in stochastic form, statistical methods of Monte Carlo Simulation (MCS) (see e.g. Young *et al.*, 1996) and Generalised Sensitivity Analysis (GSA) (Spear and Hornberger, 1980) are used to expose the poorly identifiable parameters of the model. Further, a preliminary study is made, using a DBM

combined statistical linearisation and model order reduction procedure, to determine whether a reduced, low order, model can be identified, which accurately describes the dynamics of the ecosystem.

The final chapter summaries the contributions of this thesis and suggests some possibilities for future research on the development and use of DBM modelling methods in environmental systems analysis.

CHAPTER 2

IDENTIFICATION AND ESTIMATION OF LINEAR AND NONLINEAR SYSTEMS

Over the last five decades much research has concentrated on developing methods for model identification and estimation of linear systems (see e.g. Jazwinski, 1970; Young, 1984). In contrast, because of the more complex nature of nonlinear systems, the associated techniques for system identification and estimation are in a relative stage of infancy. The main focus of the research studies presented in this chapter, centres on the development of data-based mechanistic (DBM) transfer function (TF) models for nonlinear systems and introduces the mathematical time series methods that have been utilised for their identification and estimation.

The first section of this chapter presents a brief overview of the general modelling procedure for both linear and nonlinear systems. Section 2.2 introduces the general concepts of parameter estimation. In Section 2.3 the Simplified Refined Instrumental Variable (SRIV) algorithm for parameter estimation and model identification for linear systems is discussed. Section 2.4 presents the DBM procedures for identifying

and estimating models for nonlinear systems, specifically focussing on the recursive filtering and powerful Fixed Interval Smoothing (FIS) algorithms.

2.1 GENERAL MODELLING PROCEDURE

There are four main stages in the process towards generating a DBM model of a linear or nonlinear environmental system. The first stage involves the collection of time series data that are sufficiently dynamic to enable the environmental system under consideration to be described by a TF model. For most environmental systems this stage will require the main input and output variables of the system to be identified and relevant monitoring equipment installed to record their behaviour. However, where possible, it is often extremely beneficial to perform carefully planned experiments, where the system input(s) are controlled, such that the system is perturbed over its full dynamic range.

Having obtained sufficiently informative time series data from the system, a 'suite' of model structures is selected from which the best model is chosen. The models may be: deterministic or stochastic; have a linear or nonlinear structure; have constant or time varying parameters; involve more than one input and output variable; and have different orders. The perception and experience of the modeller will determine what class of model is added to the 'suite' and will often involve additional detailed analysis of the time series data.

Having determined the 'suite' of models, the parameters of each model must be estimated. There are a great number of different algorithms for parameter estimation

but perhaps the most well known is the least squares (LS) algorithm. LS forms the basis of a number of derivative algorithms including weighted least squares (WLS), extended least squares (ELS) and generalised least squares (GLS). The favoured parameter estimation approaches adopted in the research reported in this thesis are the Simplified Refined Instrumental Variable (SRIV) and the Maximum Likelihood (ML) methods which are discussed in following sections of this chapter. Further details of these and other techniques can be found in many texts on the subject, including Jazwinski (1970), Young (1984), Ljung (1987), and Söderström and Stoica (1989).

Finally, each of the models are evaluated by objective statistical identification criteria, in order to select the 'best' overall model. When making this choice, the intended purpose of the model should not be disregarded. In some cases, the mechanistic interpretation of the optimal model structure may challenge traditional perceptions of the systems internal behaviour and require further investigation.

2.2 GENERAL PARAMETER ESTIMATION

This section introduces the general methodology underlying TF model parameter estimation. Most parameter estimation techniques are based upon the formulation of a cost function from the TF model equations which, when minimised, provides the optimal estimates of the model parameters. Consider the following generalised discrete-time, linear, TF model of a single-input-single-output (SISO) system,

$$y(k) = \frac{B(z^{-1})}{A(z^{-1})} u(k-\delta) \tag{2.1}$$

where $y(k)$ and $u(k)$ are the system output and input respectively at the kth time instant; δ is the system pure time delay; and $A(z^{-1})$ and $B(z^{-1})$ are the TF denominator and numerator polynomials respectively, defined as,

$$\begin{aligned} A(z^{-1}) &= 1 + a_1 z^{-1} + \cdots + a_n z^{-n} \\ B(z^{-1}) &= b_0 + b_1 z^{-1} + \cdots + b_m z^{-m} \end{aligned} \tag{2.2}$$

where z^{-1} is the backward shift operator, that is $z^{-i} y(k) = y(k-i)$; and the integers n and m are the number of parameters in the respective polynomials. By rearranging equation (2.1) a simple cost function can be formulated from the difference between the observed system and estimated model output, which is described as the *response error* $\hat{e}(k)$.

$$\hat{e}(k) = y(k) - \frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})} u(k-\delta) \tag{2.3}$$

However, this cost function is limited by its nonlinear parameterisation which can only be minimised by numerical methods (discussed in further detail in Section 2.5.5). An alternative arrangement of (2.1) yields the *equation error* (2.4) which is linear about its parameters and can be subsequently solved analytically, avoiding the necessity to utilise numerical techniques.

$$\hat{e}(k) = \frac{1}{\hat{A}(z^{-1})} \left[y(k)\hat{A}(z^{-1}) - \hat{B}(z^{-1}) u(k-\delta) \right] \tag{2.4}$$

One possible method of obtaining the *analytical* solution to (2.4) is to utilise the *'normal equation of linear least squares'* (see e.g. Young, 1984 for its derivation) whereby the sum of the squared errors (equation error) are analytically minimised.

The linear least squares algorithm provides good model parameter estimates, providing measured system data *without* any stochastic noise disturbance is obtained. However in practice, the observed data is usually corrupted by undesirable structured 'coloured' noise. The linear least squares algorithm acts to amplify the effects of the noise during the estimation process, further contaminating the data sequence such that parameter estimates are biased (Young, 1984). As a result, the presence of any noise on the data will cause the parameter estimates to be asymptotically biased, and statistically inconsistent regardless of the length of time series data utilised in the procedure.

2.3 SIMPLIFIED REFINED INSTRUMENTAL VARIABLE ESTIMATION AND MODEL IDENTIFICATION

To overcome this fundamental limitation of linear least squares estimation, a suite of *instrumental variable* (IV) algorithms has been developed (see e.g. Young, 1984) that provides consistent unbiased parameter estimates which require no *a priori* statistical information regarding the noise sequence. The Simplified Refined Instrumental Variable (SRIV) identification and estimation method is an extension of the original IV estimation procedure which was first introduced by Young (1970) and a simplification of the Refined Instrumental Variable (RIV) approach (see e.g. Young, 1976; Young and Jakeman, 1979; Young, 1984). It was shown (Young, 1985), that under the assumption that the noise process is a serially uncorrelated series of white noise with Gaussian distribution, that the complex RIV estimation algorithm could be reduced to the SRIV form.

2.3.1 Parameter estimation using the SRIV method

The SRIV algorithm was first introduced in the discrete-time domain but has since been extended for modelling systems using the delta operator (Young, 1991; McKenna, 1998) and in the continuous-time domain (Young and Jakeman, 1980; Foster, 1995; Price, 1999).

For the purposes of estimation, the stochastic TF model (2.1) can be conveniently rewritten in alternative vector format,

$$y(k) = \mathbf{z}(k)^{\mathrm{T}}\mathbf{a} + \eta(k); \quad k = 1,2,\cdots,N \tag{2.5}$$

where $\eta(k) = \hat{A}(z^{-1})e(k)$ and the vector $\mathbf{z}(k)$ and parameter vector $\boldsymbol{a}$ are defined in the general case respectively as,

$$\mathbf{z}(k) = [-y(k-1), -y(k-2), \cdots -y(k-n), u(k), u(k-1), \cdots, u(k-m)]^{\mathrm{T}} \tag{2.6}$$

$$\mathbf{a} = [a_1, a_1, \cdots, a_n, b_0, b_1, \cdots b_m]^{\mathrm{T}} \tag{2.7}$$

with dimensions determined by the polynomials of (2.1).

Instrumental Variable (IV) estimation involves the introduction of an *instrumental variable* (IV) vector $\hat{\mathbf{x}}(k)$ defined as follows,

$$\hat{\mathbf{x}}(k) = [-\hat{x}(k-1), -\hat{x}(k-2), \cdots -\hat{x}(k-n), u(k), u(k-1), \cdots, u(k-m)]^{\mathrm{T}} \tag{2.8}$$

Here, the *instrumental variable* $\hat{x}(k)$, is defined as the 'noise free' estimate of the system output $x(k)$ and is assumed to be serially uncorrelated with the noise process $\eta(k)$. It is generated from the *j-1*th iteration of the following 'adaptive auxiliary model',

$$\hat{x}(k) = \frac{\hat{B}_{j-1}(z^{-1})}{\hat{A}_{j-1}(z^{-1})} u(k-\delta) \tag{2.9}$$

In contrast to the simple IV method (Young, 1984), in the case of SRIV estimation, $\hat{x}(k)$ and the associated variables are 'pre-filtered' utilising the *j-1*th adaptive estimate of the auxiliary model polynomial $\hat{A}(z^{-1})$, in the following manner,

$$F_j = \frac{1}{\hat{A}_j(z^{-1})} \tag{2.10}$$

$$\hat{x}^*(k) = F_{j-1}\hat{x}(k); \quad y^*(k) = F_{j-1}y(k); \quad u^*(k) = F_{j-1}u(k) \tag{2.11}$$

The adaptive pre-filter F_j acts to remove any undesirable high frequency components from the input signal, which would otherwise reduce the efficiency of the estimation results, whilst retaining those frequencies that are essential for system analysis.

For a given sample size *N*, the non-recursive optimal SRIV estimate $\hat{\boldsymbol{a}}$ of the parameter vector $\boldsymbol{a}$ is then obtained from the solution of the following SRIV algorithm,

$$\left[\sum_{k=1}^{k=N} \hat{\boldsymbol{x}}^*(k)\boldsymbol{z}^*(k)^{\mathrm{T}}\right]\hat{\boldsymbol{a}} = \sum_{k=1}^{k=N} \hat{\boldsymbol{x}}^*(k)\, y^*(k) \tag{2.12}$$

where,

$$\mathbf{z}^*(k) = \left[-y^*(k-1), -y^*(k-2), \cdots, -y(k-n)^*, u(k)^*, u^*(k-1), \cdots, u^*(k-m)\right]^{\mathrm{T}} \tag{2.13}$$

$$\hat{\mathbf{x}}^*(k) = \left[-\hat{x}^*(k-1), \cdots -\hat{x}^*(k-n), u^*(k), u^*(k-1), \cdots u^*(k-m)\right]^{\mathrm{T}} \tag{2.14}$$

In the first instance, to initiate the algorithm the *a priori* estimate of the IV vector $\hat{\boldsymbol{x}}^*(k)$ can be obtained from the 'auxiliary model' using linear least squares parameter estimates. This *preliminary* IV vector is subsequently inserted into the SRIV algorithm to yield an *initial* estimate of the parameter vector $\hat{\boldsymbol{a}}$. The statistical efficiency of $\hat{\boldsymbol{a}}$ is further refined through subsequent iterations of this procedure, where at each iteration, the parameters of the adaptive pre-filter and auxiliary model are updated. It has been demonstrated, however (Young, 1976, 1985; Young and Jakeman, 1979), that optimum SRIV parameter estimates can be obtained following only three iteration steps.

Having determined the optimal parameter vector $\hat{\boldsymbol{a}}$ from the final iteration step, the SRIV algorithm also generates the *inverse of the instrumental product matrix* (IPM) $\hat{\boldsymbol{P}}(N)$ defined as,

$$\hat{\mathbf{P}}(N) = \left[\sum_{k=1}^{k=N} \hat{\mathbf{x}}^*(k)\hat{\mathbf{x}}^*(k)^{\mathrm{T}}\right]^{-1} \tag{2.15}$$

and the *covariance matrix* $\boldsymbol{P}^*(N)$,

$$\boldsymbol{P}^*(N) = \hat{\sigma}^2\hat{\boldsymbol{P}}(N) \tag{2.16}$$

where an estimate $\hat{\sigma}^2$ of the noise variance σ^2 can be obtained from model residuals $\hat{e}(k)$ where

$$\hat{e}(k) = y(k) - \hat{x}(k); \text{ and } \hat{\sigma}^2 = \frac{1}{N}\sum_{k=1}^{k=N}\hat{e}(k)^2 \tag{2.17}$$

The covariance matrix $\boldsymbol{P}^*(N)$ provides essential information regarding the level of uncertainty associated with each parameter estimate. The leading diagonal elements of the matrix distinguish the *variance* of each parameter estimate whilst a measure of the parameter *covariance* is provided by the matrix off diagonal elements. In this regard, the covariance matrix can be utilised within the context of Monte Carlo uncertainty and sensitivity analyses (see Chapters 4 and 6).

2.3.2 Model Order Identification

Having estimated the parameters of a variety of different models, an optimal model structure must be selected. Model order identification, namely the process of identifying the most appropriate values of n, m and δ in (2.1), is chiefly undertaken, although not exclusively, with the assistance of carefully selected objective statistical criteria. In combination, these objective methods should provide both a measure of how well the model output explains the data and indicate the presence of model over-parameterisation. In this study, model identification is based upon the coefficient of determination, R_T^2, and Young's Identification Criterion, YIC, (see e.g. Young, 1989). The coefficient of determination provides a statistical measure of how close the model output fits the observed system output and is defined as,

$$R_T^2 = 1 - \frac{\hat{\sigma}^2}{\hat{\sigma}_y^2} \tag{2.18}$$

where $\hat{\sigma}^2$ is the sample variance of the model residuals $\hat{e}(k)$ (2.3), and $\hat{\sigma}_y^2$ is the sample variance of the measured output $y(k)$ about its mean value $\bar{y}$. Clearly, a good model fit is obtained where the variance of the model residuals $\hat{\sigma}^2$ is low in comparison to the variance of the data $\hat{\sigma}_y^2$ and the R_T^2 approaches unity. Conversely, a poor model has a residual variance $\hat{\sigma}^2$ that is close to the magnitude of the sample variance $\hat{\sigma}_y^2$ and the R_T^2 tends towards zero. It is important to differentiate R_T^2 from the more conventionally adopted coefficient of determination R^2. The latter is based upon the variance of the one-step ahead prediction errors, rather than the model response errors and, whilst this is a popular criterion for assessing the performance of forecasting models, it is less suited for TF model order identification. In particular, model *one-step ahead prediction errors* are relatively straightforward to minimise, as the predictions are based upon past values of the system output itself; in contrast, the model *response errors* are more difficult to minimise as the model output is formulated based on the system input only. In a hydrological context, R_T^2 is also commonly called the Nash and Sutcliffe efficiency criterion with unity power.

Although an excellent measure of model fit, the R_T^2 criterion should not generally be utilised independently to assess the merits of a model, since it does not consider the models relative complexity or the level of uncertainty associated with the parameter estimates.

For this reason, an additional statistical measure, which incorporates both of these features, is utilised within the process of model identification. The heuristic Young's Identification Criterion YIC, is defined as

$$YIC = \log_e \frac{\hat{\sigma}^2}{\hat{\sigma}_y^2} + \log_e \{NEVN\}; \quad NEVN = \frac{1}{np} \sum_{i=1}^{i=np} \frac{\hat{\sigma}^2 \hat{p}_{ii}}{\hat{a}_i^{\,2}} \tag{2.19}$$

where the variables of the leading term are defined as in equation (2.18); $np = n + m + 1$ is the number of parameters in the TF model (2.1) denominator (n) and numerator (m) polynomials; $\hat{a}_i$ is the ith element in the parameter vector $\hat{\boldsymbol{a}}(N)$; $\hat{p}_{ii}$ is the ith diagonal element of the covariance matrix $\hat{\mathbf{P}}(N)$ (where N is the total number of samples); such that $\hat{\sigma}^2 \hat{p}_{ii}$ is an estimate of the error variance associated with the ith element of the parameter vector $\hat{\boldsymbol{a}}(N)$ after N samples.

The first term of equation (2.19) is a normalised measure of how well the model fits the data; as the model fit improves, the ratio of $\hat{\sigma}^2/\hat{\sigma}_y^2$ decreases and the term becomes more negative. The second term provides an indication of the relative uncertainty of each ith parameter estimate, normalised with respect to all the parameters in the npth order model; as the parameter error variance decreases, so the parameter is better defined and the second term becomes smaller. A model that fits the data well, but has a high order, with ill defined parameters, will have a correspondingly high YIC value because of the large magnitude of the second term. The YIC criterion, therefore, provides a compromise between model fit and over parameterisation. An optimal model, with a low YIC, should give a good fit to the data and have well defined parameters.

In practice, the minimisation of the *YIC* will not necessarily identify the best overall model and should, therefore, be used in conjunction with the R_T^2 criterion. This will ensure that the model selected will fit the data well, without compromising parametric efficiency and uncertainty. Additionally, it is important not to carry out the process of model order identification without due regard to the physical nature of the system under consideration. For example, if the system is known to have parallel or feedback processes operating within it, model structures that can describe these behaviours should be evaluated. Furthermore, the philosophy underpinning DBM modelling should not be disregarded during the identification stage; a model structure that has a clear physical interpretation may be favoured for selection over an alternative structured model with slightly superior identification statistics. An additional, tertiary model order identification criterion utilised in the research reported in this thesis was the AIC criterion (see Box and Jenkins, 1970).

2.4 DBM IDENTIFICATION AND ESTIMATION OF NONLINEAR SYSTEMS

Many environmental systems are non-stationary and nonlinear in nature and as a consequence, alternative approaches to model identification and estimation need to be adopted in order to develop models that successfully characterise their behaviour, e.g. the EKF (Whitehead and Hornberger, 1982; Chen and Beck, 1993). In this section a novel DBM approach to modelling nonlinear systems is presented which has been successfully applied in the areas of economics, ecology, biology, engineering and environmental science (Young, 1993; Young and Beven, 1994; Young and Pedregal,

1997; Young, 1998a). The DBM approach has the overall objective of identifying a nonlinear TF with time invariant parameters through a process of objective statistical inference applied to the time series data.

The preliminary stage of the DBM methodology is to determine that the system in consideration is in fact nonlinear. This can be ascertained in the first instance, by analysing the residuals of the best SRIV identified constant parameter linear TF model, using standard statistical tests (e.g. Billings and Voon, 1986). In the second instance, this may be achieved by allowing the parameters of a linear TF to vary over time. The powerful Fixed Interval Smoothing (FIS) method of recursive estimation (Young, 1993) can be used to obtain a non-parametric estimate of the model time varying parameters (TVP). Any parameter that is found to be significantly time variant over the observation interval, may reflect the non-stationary or nonlinear behaviour of the system. Analysis then proceeds to investigate whether the identified temporal parameter variations are state dependent and can be efficiently parameterised. In accordance with DBM philosophy, in addition to enhancing the model fit to the data, it is essential that the identified state has a clear mechanistic interpretation in relation to the system under consideration. Having identified the structure of the nonlinear dynamic model, the final stage of the DBM methodology is to re-estimate all the model parameters against the time series through nonlinear optimisation.

2.4.1 Background

The following discussion assumes that the behaviour of a discrete-time, dynamic nonlinear and/or non-stationary environmental system can be represented by a general dynamic, stochastic model which can be expressed as;

$$y(k) = \Im\{\boldsymbol{\chi}(k), \mu(k)\} \tag{2.20}$$

where $y(k)$ is the measured output of the system under consideration and $\Im\{\cdot\}$ can be regarded as a well behaved nonlinear function of the variables that are contained within the non minimum state space (NMSS) state vector $\boldsymbol{\chi}(k)$,

$$\boldsymbol{\chi}(k) = \left[y(k-1), \cdots, y(k-n)\boldsymbol{u}(k)^{\mathrm{T}}, \cdots, \boldsymbol{u}(k-m)^{\mathrm{T}}\boldsymbol{U}(k)^{\mathrm{T}}, \cdots, \boldsymbol{U}(k-q)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{2.21}$$

Here, $\boldsymbol{\chi}(k)$ constitutes a vector of past values of the system output $y(k)$; present and past values of the deterministic input variables $\boldsymbol{u}(k)$ that are considered to be the *principle* inputs to the system; and vector $\boldsymbol{U}(k)$, which contains present and past values of additional input variables that *may* affect the system nonlinearly, which at this preliminary stage of the analysis, are yet to be determined. In addition, the unobserved, serially uncorrelated, zero mean white noise process $\mu(k)$, introduces the stochasticity into the model and is assumed to be statistically independent to the input variables $\boldsymbol{u}(k)$ and $\boldsymbol{U}(k)$.

For explanatory purposes, consider a nonlinear system with only *one* principle input variable. By adopting the theory of model linearisation (see e.g. Young, 1993), it is

reasonable to assume that the nonlinear model (2.20) can, in most cases, be approximated by a linear TF with *time varying parameters,* which in the general form can be written as,

$$y(k) = \frac{b_0(k) + b_1(k)z^{-1} +, \cdots, + b_m(k)z^{-m}}{1 + a_1(k)z^{-1} +, \cdots, + a_n(k)z^{-n}} u(k) + \zeta(k) \tag{2.22}$$

Here, the time variant nature of the parameters may reveal any non-stationary or nonlinear behaviour present in the system. The noise term $\zeta(k)$ arises from stochastic disturbances in $\mu(k)$.

2.4.2 Time varying parameter (TVP) estimation using the Fixed Interval Smoothing (FIS) algorithm

Having presented the theoretical background to the DBM nonlinear modelling concept, the following section discusses how the time variable parameters (TVP) are estimated using a two pass recursive (forward filtering and backward smoothing) operation on the time series data.

Pass 1: 'Forward filtering'

In vector format, equation (2.22) can be represented in the form,

$$y(k) = \boldsymbol{z}(k)^{\mathrm{T}} \boldsymbol{a}(k) + \eta(k); k = 1,2,\cdots,N \tag{2.23}$$

where the state and parameter vectors are now defined as,

$$z(k) = [-y(k-1), -y(k-2), \cdots -y(k-n), u(k), u(k-1), \cdots u(k-m)]^T \tag{2.24}$$

$$\boldsymbol{a}(k) = [a_1(k), a_2(k), \cdots a_n(k), b_0(k), b_1(k), \cdots b_m(k)]^T \tag{2.25}$$

and where the noise process $\eta(k)$ is assumed initially to be 'white' noise with variance σ^2. Having formulated model (2.21), it is necessary to characterise the temporal variation shown by the parameters in vector $\boldsymbol{a}(k)$ by some form of mathematical description. One method, that has been demonstrated by Young (1978, 1984, 1993) as capable of modelling these parameter variations, is the following Gauss-Markov stochastic process,

$$\mathbf{x}(k) = \mathbf{F}(k)\mathbf{x}(k-1) + \mathbf{G}(k)\mathbf{q}(k) \tag{2.26}$$

where $\mathbf{x}(k)$ is a 'state vector' representing the parameters in $\boldsymbol{a}(k)$ as well as other elements that are necessary to characterise their complete temporal evolution. $\boldsymbol{F}(k)$ and $\boldsymbol{G}(k)$ are transition and input matrices respectively and $\boldsymbol{q}(k)$ is a vector of serially uncorrelated, random noise with zero mean and covariance $\boldsymbol{Q}$. Here, the transition matrix $\boldsymbol{F}(k)$ determines the relationship between successive parameter vectors, whilst the magnitude of parameter variation, introduced into the model by the stochastic disturbance $\boldsymbol{q}(k)$, is regulated by the input matrix $\boldsymbol{G}(k)$ and $\boldsymbol{Q}$. It is worth noting, that when $\boldsymbol{F}(k)$ and $\boldsymbol{G}(k)$ are both identity matrices, equation (2.24) can be simplified to the well known *random walk* (RW) model,

$$\mathbf{x}(k) = \mathbf{x}(k-1) + \mathbf{q}(k) \tag{2.27}$$

It follows that from (2.26) that the best *a priori* estimate of vector $\hat{\mathbf{x}}(k)$, at the previous time instant, is generated simply using,

$$\hat{\mathbf{x}}(k \mid k-1) = \mathbf{F}\hat{\mathbf{x}}(k-1) \tag{2.28}$$

where the argument $(k|k-1)$ represents the estimate of $\hat{\mathbf{x}}$ at time *k*, conditional on the previous estimate at *k*-1. The *a priori* estimate of the covariance matrix associated with equation (2.28) is obtained from,

$$\hat{\mathbf{P}}(k \mid k-1) = \mathbf{F}(k)\hat{\mathbf{P}}(k-1)\mathbf{F}(k)^{\mathrm{T}} + \mathbf{G}(k)\mathbf{Q}_r\mathbf{G}(k)^{\mathrm{T}} \tag{2.29}$$

It has been shown (Young, 1984) that for estimation purposes, the ratio of $\boldsymbol{Q}$ and σ^2 is important, rather than their explicit values. Therefore, the *noise variance ratio* (NVR) matrix $\boldsymbol{Q}_r$ is introduced into the estimation procedure and is defined as

$$\boldsymbol{Q}_r = \boldsymbol{Q}/\sigma^2 \tag{2.30}$$

If model (2.23) was defined in its alternative state space form where the state vector $\mathbf{x}(k) = \mathbf{a}(k)$ and the observation vector $\mathbf{H}(k) = \boldsymbol{z}(k)^{\mathrm{T}}$, a recursive, prediction-correction filtering algorithm for estimating the time varying parameters in model (2.23) can be formed by combining equations (2.28 - 2.29) with a *recursive* time variant form of the least squares parameter estimation algorithm. Thus, the 'forward pass' of the TVP estimation routine is defined and takes the following form,

Prediction

$$\begin{aligned} \hat{\mathbf{x}}(k \mid k-1) &= \mathbf{F}(k)\hat{\mathbf{x}}(k-1) \\ \hat{\mathbf{P}}(k \mid k-1) &= \mathbf{F}(k)\hat{\mathbf{P}}(k-1)\mathbf{F}(k)^{\mathrm{T}} + \mathbf{G}(k)\mathbf{Q}_r\mathbf{G}(k)^{\mathrm{T}} \end{aligned} \tag{2.31}$$

Correction

$$\begin{aligned}\hat{\mathbf{x}}(k) &= \hat{\mathbf{x}}(k-1) + \hat{\mathbf{P}}(k \mid k-1)\mathbf{H}(k)^{\mathrm{T}} \\ &\quad \times \left[1 + \mathbf{H}(k)\hat{\mathbf{P}}(k \mid k-1)\mathbf{H}(k)^{\mathrm{T}}\right]^{-1}\left[y(k) - \mathbf{H}(k)\hat{\mathbf{x}}(k \mid k-1)\right] \\ \hat{\mathbf{P}}(k) &= \hat{\mathbf{P}}(k \mid k-1) - \hat{\mathbf{P}}(k \mid k-1)\mathbf{H}(k)^{\mathrm{T}} \\ &\quad \times \left[1 + \mathbf{H}(k)\hat{\mathbf{P}}(k \mid k-1)\mathbf{H}(k)^{\mathrm{T}}\right]^{-1}\mathbf{H}(k)\hat{\mathbf{P}}(k \mid k-1)\end{aligned} \tag{2.32}$$

A priori predictions of the parameter vector are, therefore, based on past knowledge of the state (parameter) vector variation. As the algorithm continues to pass through the data, these *a priori* predictions are corrected by (2.32) to yield *a posteriori* estimates in a recursive fashion.

The recursive filtering algorithm (equations 2.31-2.32) is identical in form to the well known Kalman filter (Kalman, 1960). However, because the observation vector $\mathbf{H}(k)$ is composed of stochastic variables contaminated by measurement noise rather than by purely deterministic variables, as necessitated by the formal definition of the Kalman Filter, algorithm (2.31-2.32) is described here in a recursive least squares context.

NVR estimation

The NVR matrix $\boldsymbol{Q}_r$, which is normally chosen to be diagonal, with NVR elements $q_{ii}, i = 1,2,\cdots,m+n+1,$ is crucial in the overall TVP estimation process as it controls the magnitude of model parameter variation. The NVR values in $\boldsymbol{Q}_r$ may fundamentally affect the success of any subsequent state dependent analysis and, therefore, the overall ability to characterise any nonlinear behaviour in the data. If an NVR is chosen too large, the associated parameter estimate will change rapidly and

the model will subsequently fit the data very well. However, this is unlikely to reveal any information regarding a physically meaningful trend in parameter variation. Alternatively, if a very small NVR is selected, the parameter estimates may vary too slowly to expose a trend that can be tested for state dependence.

Maximum Likelihood (ML) optimisation is an objective method, with a strong theoretical background, that can be utilised to estimate an optimal NVR, in order to reveal the best trend in the TVP's. Estimation of the NVR via ML was adopted in the analyses reported in this thesis. However, this method occasionally fails with some data series, due to the flat nature of the likelihood surface that is often synonymous with ML within the present setting. On these occasions, the NVR has been carefully selected manually such to provide parameter estimates that are visually acceptable in the present context. The theoretical background to ML will be discussed in Section 2.5.2.

Pass 2: 'Backward Smoothing'

The 'forward pass' of the TVP estimation procedure provides estimates of $\mathbf{x}(k)$ based upon the known behaviour of the parameter variation up to, and including, the kth sampling instant, e.g. $\hat{\mathbf{x}}(k \mid k)$. It is usual, when estimating off-line, that the complete data series are available for analysis. Under these normal circumstances, the initial estimate of the TVP vector $\hat{\mathbf{x}}(k)$ can be re-estimated at the kth sampling instant, conditional on the data series over the *whole* observational interval (e.g. $k = 1, 2, \cdots, N$), to provide a refined or 'smoothed' estimate $\hat{\mathbf{x}}(k \mid N)$. This beneficial backward recursive pass of the data, 'smoothes' the estimate of the parameter vector

and reduces the variance of the estimated model error. An additional advantage of this second step, is that it ensures that $\hat{\mathbf{x}}(k \mid N)$ does not suffer from 'phase lag' which is an intrinsic problem of the filtered estimate $\hat{\mathbf{x}}(k \mid k)$ (Young, 1993). There are several *fixed interval smoothing* algorithms available (see e.g. Norton, 1975), but the one utilised in the research reported in this thesis, takes the following form,

$$\hat{\mathbf{x}}(k \mid N) = \mathbf{F}(k)\left[\hat{\mathbf{x}}(k+1 \mid N) + \mathbf{G}(k)\mathbf{Q}_r\mathbf{G}(k)^{\mathrm{T}}\mathbf{L}(k-1)\right] \tag{2.33}$$

where the recursive estimate of the 'Lagrange multiplier' vector $\boldsymbol{L}(k)$ is obtained from,

$$\begin{aligned} \boldsymbol{L}(N) &= 0 \\ \mathbf{L}(k-1) &= \left[\mathbf{I} - \mathbf{P}(k)\mathbf{H}(k)^{\mathrm{T}}\mathbf{H}(k)\right] \\ &\quad \times \left\{\mathbf{F}(k)^{\mathrm{T}}\mathbf{L}(k) - \mathbf{H}(k)^{\mathrm{T}}\left[y(k) - \mathbf{H}(k)\mathbf{F}(k-1)\hat{\mathbf{x}}(k-1)\right]\right\} \end{aligned} \tag{2.34}$$

In theory, the vector $\mathbf{H}(k)$ should contain only pure deterministic variables. However, in low noise situations, this vector is replaced 'sub-optimally' to contain stochastic variables contaminated by low measurement noise. Recently, this restriction has been removed by the development of an IV smoothing algorithm (Young and McKenna, 1999). However, this has not been employed in the present research studies.

An advantage of algorithm (2.33-2.34) over other fixed interval smoothing methods, (see e.g. Priestly, 1988) is that, under the assumption that the estimated model error $\hat{\eta}(k)$ (equation 2.23) is a sequence of serially uncorrelated random variables with zero mean, the covariance matrix $\boldsymbol{P}^*(k|N)$ of the smoothed parameter estimates can be

obtained directly (Young, 1993). The smoothed estimate of the covariance matrix is defined as,

$$\boldsymbol{P}^*(k|N) = \hat{\sigma}^2 \boldsymbol{P}(k|N) \tag{2.35}$$

$$\begin{aligned} \mathbf{P}(k \mid N) = \mathbf{P}(k) + \mathbf{P}(k)\mathbf{F}(k+1)^{\mathrm{T}}\left[\mathbf{P}(k+1 \mid k)\right]^{-1} \\ \times\left[\mathbf{P}(k+1 \mid N) - \mathbf{P}(k+1 \mid k)\right]\left[\mathbf{P}(k+1 \mid k)\right]^{-1}\mathbf{F}(k+1)\mathbf{P}(k) \end{aligned} \tag{2.36}$$

In circumstances where the estimated model error $\hat{\eta}(k)$ is not Gaussian white noise, the covariance matrix $\boldsymbol{P}^*(k|N)$ provides useful information about the uncertainty of each parameter estimate. Once again, the recently developed IV fixed interval smoothing algorithm efficiently removes the restriction of Gaussian white noise.

Having estimated $\hat{\mathbf{x}}(k \mid N)$, the magnitude of each parameter variation can be quantified. If the TVP estimates appear relatively stationary over the observation interval, the system is principally linear and should be well approximated by a linear TF model. However, if any parameter exhibits significant temporal variation, further analyses to investigate possible state dependence should ensue. The preceding analyses should be performed utilising the simplest model that is capable of representing the system behaviour (for example a 1st order TF). Subsequent model identification at the final stage of the analyses will indicate whether it is necessary to increase the model order.

2.4.3 State dependent parameter modelling

The preceding section has discussed a method for effectively linearising model (2.20) by obtaining a non-parametric estimate of the TVP vector $\mathbf{a}(k)$ (equation 2.23) through a two step recursive procedure. A forward 'filtering' pass of the time series provides an initial estimate of $\hat{\boldsymbol{a}}(k|k)$ which is then 'smoothed' through a backward pass of the data to obtain the final estimate $\hat{\boldsymbol{a}}(k|N)$.

The analysis subsequently proceeds to investigate whether the identified temporal parameter variations, associated with the nonlinear behaviour of the system, can be explained by nonlinear functions of the other state variables present within the system NMSS vector $\boldsymbol{\chi}(k)$ which are yet to be incorporated into the model.

One possible approach to investigating state dependency is to assume that $\boldsymbol{a}(k)$ exhibits a linear relation to nonlinear functions of $\boldsymbol{\chi}(k)$, such that,

$$\boldsymbol{a}(k) = \boldsymbol{M}\{\boldsymbol{\chi}(k)\}\boldsymbol{\alpha}(k) \tag{2.37}$$

where $\boldsymbol{M}(k) = \boldsymbol{M}\{\boldsymbol{\chi}(k)\}$ constitutes an n by m transformation matrix of nonlinear functions, dependent on appropriately selected state variables from the NMSS vector $\boldsymbol{\chi}(k)$ and $\boldsymbol{\alpha}(k)$ is an m-dimensional parameter vector. In ideal circumstances, the appropriate choice of state variables and nonlinear functions will cause parameter vector $\boldsymbol{\alpha}(k)$ to be time invariant. However, some applications (e.g. adaptive control and forecasting) may actually require the parameter vector $\boldsymbol{\alpha}(k)$ to be time variant. The state dependent parameter modelling (SDPM) presented in this thesis did not

necessitate the elements in $\alpha(k)$ to be temporally variant and therefore, specific methods of estimation in this context will not be discussed. However, for completeness, it is sufficient to say, that the temporal variation of $\alpha(k)$ can be modelled as a Gauss-Markov 'random walk' process, e.g.

$$\alpha(k) = \alpha(k-1) + \upsilon(k) \tag{2.38}$$

where $\upsilon(k)$ is a serially uncorrelated vector of random variables with zero mean. The reader is encouraged to consult Young (1993) for further information on this subject.

The identification of appropriate nonlinear transformed state variables from $\boldsymbol{M}(k)$ is fundamental to the success of the SDPM process. Various methods can be utilised to assist in this selection process, such as scatter plots and correlation analysis in which the elements of $\boldsymbol{M}(k)$ are compared to the TVP vector $\hat{\boldsymbol{a}}(k|N)$. It is important that the uncertainty associated with the estimates of $\hat{\boldsymbol{a}}(k|N)$, the covariance matrix $\boldsymbol{P}^*(k|N)$, is incorporated within the selection process. Moreover, it must be strongly emphasised, that in line with the DBM modelling philosophy, the state variables and nonlinear functions chosen, must have some realistic physical interpretation in relation to the nature of the system under consideration.

On the basis that α has constant parameters, $\boldsymbol{M}(k)$ can be determined in the following manner,

$$\hat{\boldsymbol{a}}(k|N) = \boldsymbol{M}(k)\alpha + \varepsilon(k) \tag{2.39}$$

where $\hat{\boldsymbol{a}}(k|N)$ is derived from the FIS algorithm and the noise vector $\boldsymbol{\varepsilon}(k)$ has covariance $\boldsymbol{P}^*(k|N)$. Exploiting knowledge of the $\boldsymbol{P}^*(k|N)$ matrix, $\boldsymbol{\alpha}$ can be estimated by minimising the following weighted least squares (WLS) cost function,

$$J = \sum_{k=1}^{k=N} [\hat{\boldsymbol{a}}(k|N) - \boldsymbol{M}(k)\boldsymbol{\alpha}]^{\mathrm{T}} \boldsymbol{W}(k) [\hat{\boldsymbol{a}}(k|N) - \boldsymbol{M}(k)\boldsymbol{\alpha}] \tag{2.40}$$

where $\boldsymbol{W}(k) = \boldsymbol{P}^*(k|N)^{-1}$ is normally based on only the diagonal elements of $\boldsymbol{P}^*(k|N)$. The introduction of $\boldsymbol{P}^*(k|N)$ into algorithm (2.40), ensures that in estimating $\boldsymbol{\alpha}$, a changing estimate of the uncertainty associated with the parameter vector $\hat{\boldsymbol{a}}(k|N)$ is incorporated. Alternative methods of estimation are discussed in Young (1998a).

It has been demonstrated that the process of state dependent parameter modelling can sometimes be enhanced by sorting the data in some non-temporal order (with respect to either the input or output series) prior to the estimation of model time varying parameters. In this manner, the rapid natural variations between the input and output time series are effectively eliminated in the 'sorted' data space such that the estimated time varying parameters are much smoother, exhibiting less rapid variations. The process of state dependent parameter modelling can then proceed, investigating the relationship between the 'smoother' estimated time varying parameters with appropriate variables taken from the NMSS vector (sorted in an identical manner). This procedure has led to the identification of better defined state dependent parameter relationships.

2.4.4 Illustrative example: The logistic growth equation.

To demonstrate the efficacy of the DBM modelling approach to the identification of nonlinear systems, consider the following illustrative example. Equation (2.41) presents an example of the logistic growth equation; a simple model of population growth. Logistic growth is characterised by an initial exponential behaviour being limited by a ceiling or capacity.

$$y(k) = 4y(k-1) - 4y(k-1)^2 \tag{2.41}$$

Figure 2.1 presents the simulated response of the logistic growth equation which was initiated (k=1) with a value of 0.1. A small amount of measurement noise has been added to this data to make the example more realistic.

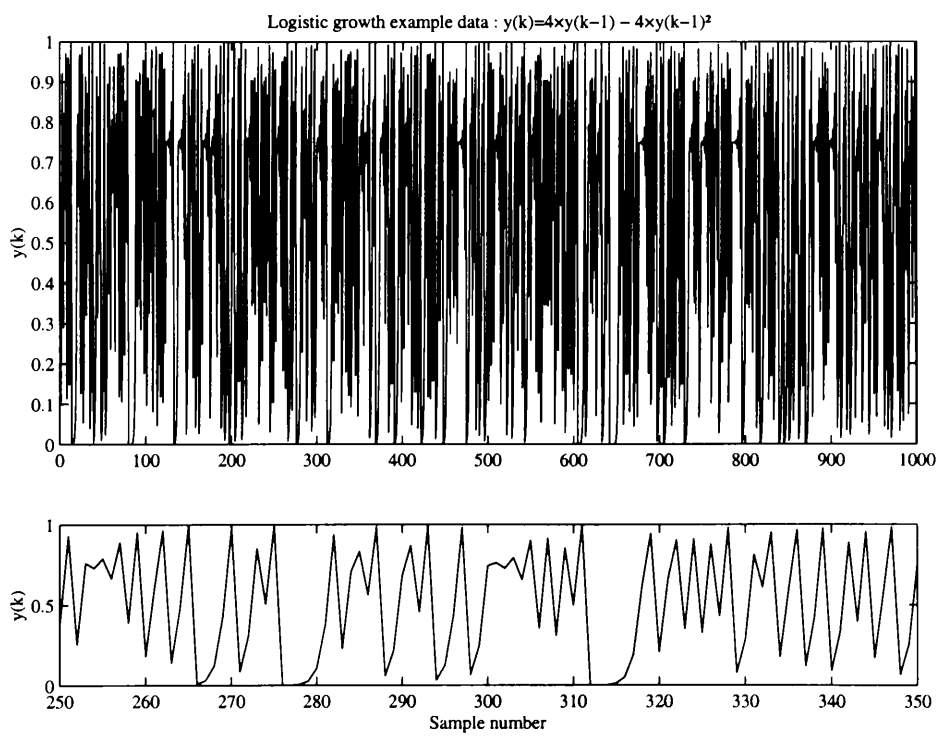

Figure 2.1 Response of the logistic growth equation.

The FIS algorithm can be utilised to provide an estimate $\hat{a}(k \mid N)$ of the TVP $a(k)$ which characterises the nonlinear behaviour between the input $y(k-1)$ and output $y(k)$ series as shown by equation (2.42).

$$y(k) = a(k)y(k-1) \tag{2.42}$$

Prior to estimation, the input data series $y(k-1)$ was sorted in ascending order of numerical magnitude (using the sort method in MATLAB) and the output series $y(k)$ was reordered in the same manner. Figure 2.2 presents the FIS estimate $\hat{a}(\circ \mid N)$ of the TVP based on input and output data sorted in a non-temporal order.

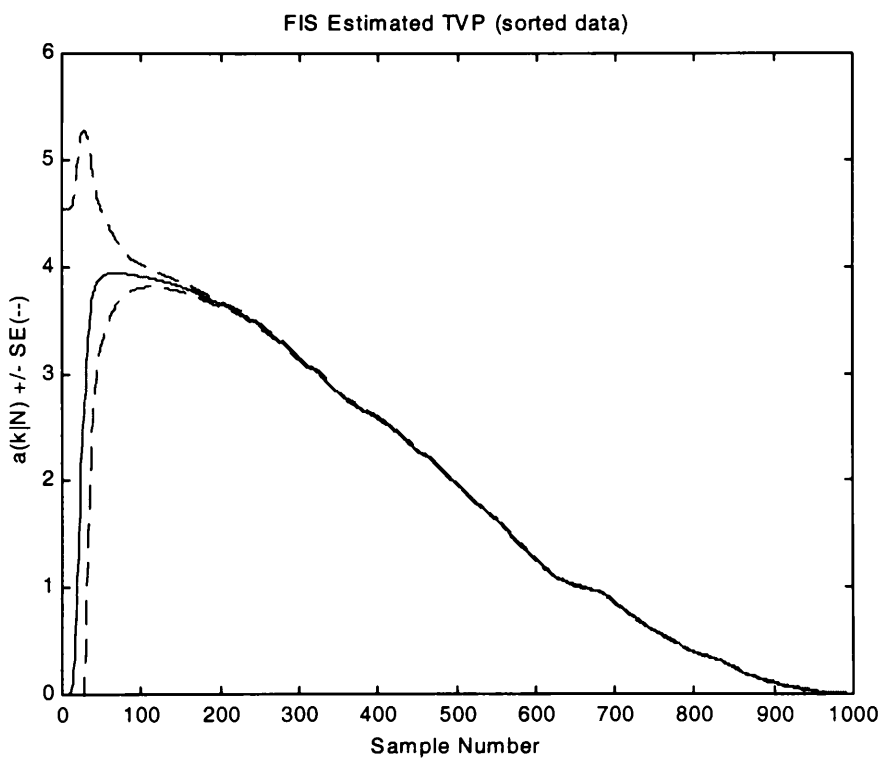

Figure 2.2 FIS estimate $\hat{a}(\circ \mid N)$ with standard errors shown as a dashed lines.

In normal circumstances, where there is no prior knowledge of the system nonlinearity, the TVP would be compared against other variables from the NMSS vector. However, in this simulated example, the nature of the system nonlinearity is already known ($a(k) = 4 - 4y(k-1)$). If the TVP estimate $\hat{a}(\circ \mid N)$ has managed to capture this nonlinear behaviour, a plot of $\hat{a}(\circ \mid N)$ versus $y(k-1)$ should yield a straight line with an intercept of 4 and a gradient of –4. Figure 2.3 confirms the success of the TVP procedure: $\hat{a}(\circ \mid N)$ versus $y(k-1)$ is plotted as a full line with standard error bounds plotted as dashed lines. Equation $4 - 4y(k-1)$ is also plotted. The comparison between the lines is very good, with only a slight deviation observed at low values of $y(k-1)$.

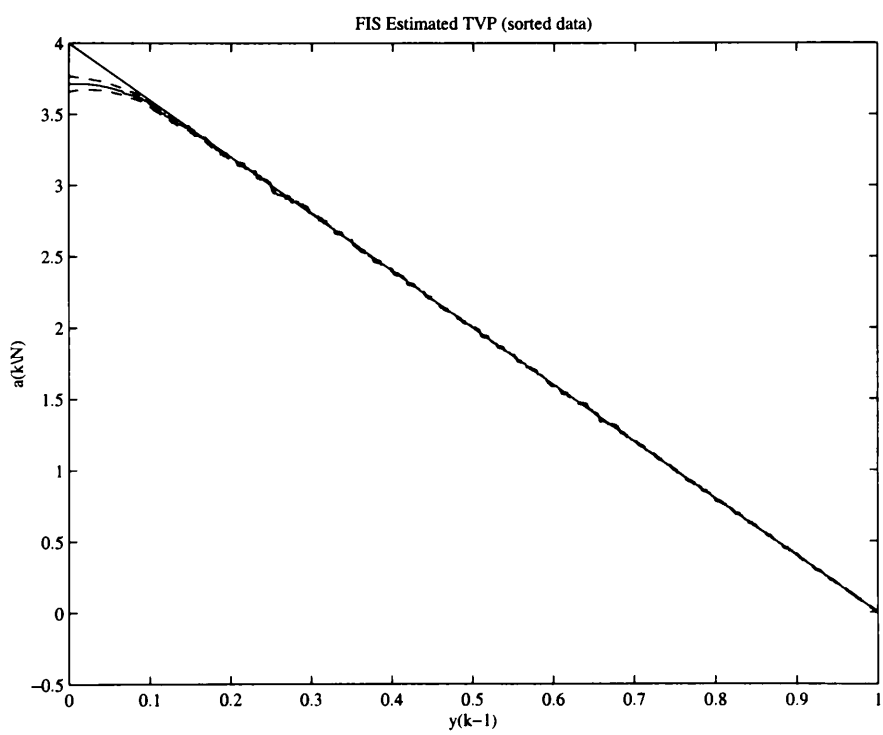

Figure 2.3 FIS estimate $\hat{a}(\circ \mid N)$ versus $y(k-1)$ with standard errors (red line). Actual nonlinear relationship $4 - 4y(k-1)$ (blue line).

Whilst this example has been idealised by using simulated data, it has highlighted how effective the DBM techniques are at identifying the nonlinearity between related time series.

2.4.5 Final estimation of a nonlinear model

Having identified that the system under consideration is nonlinear, the FIS algorithm has provided an estimate of the time varying parameter vector $\hat{\boldsymbol{a}}(k|N)$ which characterises the system nonlinearity. The SDPM modelling procedure outlined above has identified the *structure* of the final nonlinear model, but it does not provide statistically efficient estimates of the model parameters. In the final stage of the analyses, all the parameters of the nonlinear model are re-estimated. Deterministic or stochastic methods of optimisation can be utilised to obtain the final model parameter estimates independently or in conjunction with the SRIV algorithm. A brief overview of additional numerical optimisation methods adopted for the research reported in this thesis, are given in Section 2.5. Figure 2.4 summarises, in schematic form, the complete DBM modelling procedures for nonlinear identification and estimation.

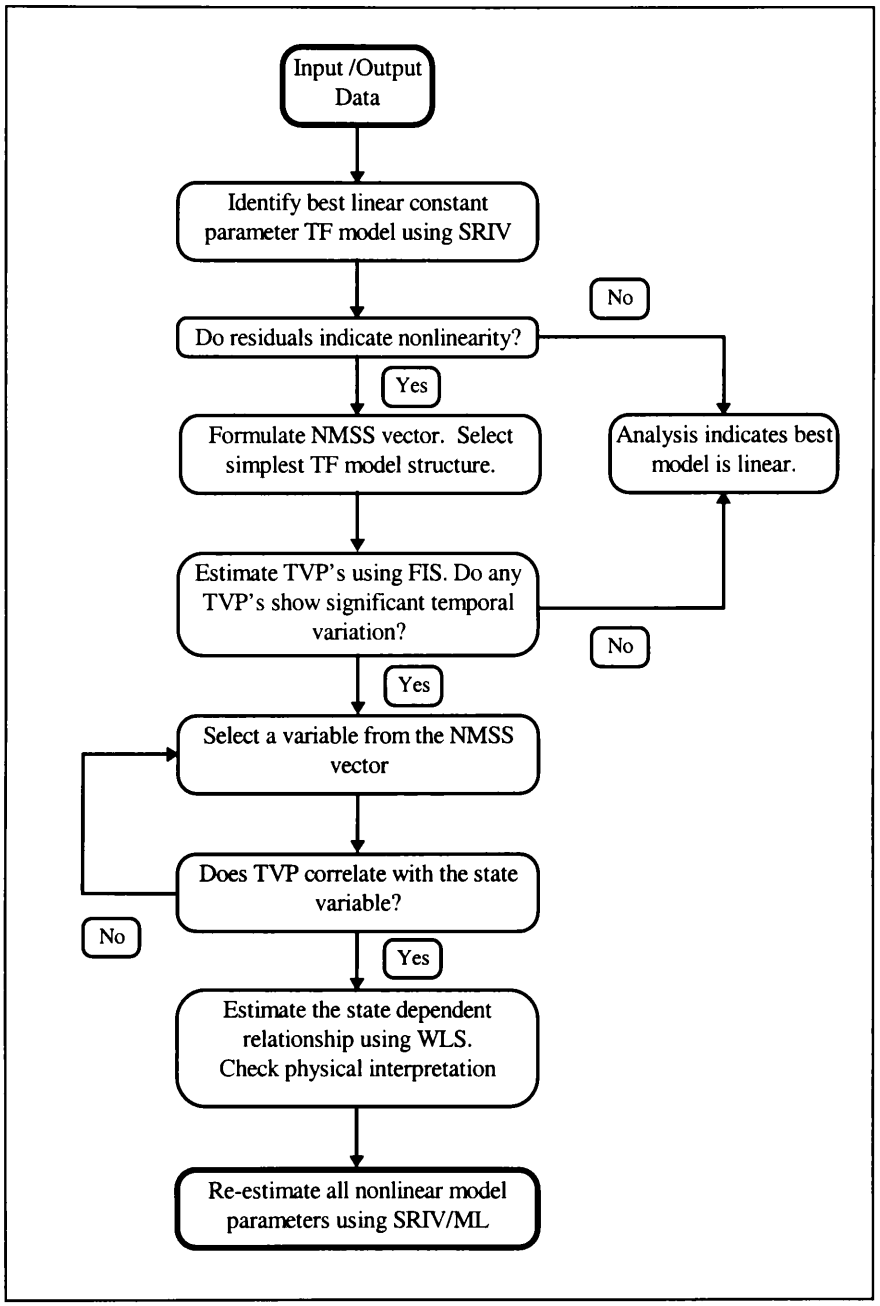

Figure 2.4 A summary of the complete DBM modelling procedure

2.5 STOCHASTIC AND DETERMINISTIC METHODS OF PARAMETER ESTIMATION

This section briefly discusses two contrasting numerical approaches for estimating the parameters of nonlinear models. These techniques can be applied to nonlinear models (and indeed linear models) whose structures have been determined either through statistically objective identification procedures, e.g. DBM identification (Section 2.3), or by alternative deterministic approaches.

For nonlinear models, the estimation problem is two fold. Firstly, an appropriate cost function must be selected to provide a measure of the difference between the model and system outputs. This can either be defined deterministically, based upon the least squares method, or stochastically, based upon the more objective ML approach (see e.g. Harvey, 1989). In contrast to linear models, cost functions cannot be solved analytically to provide optimal model parameter estimates and so it is necessary to select a suitable numerical technique to optimise (minimise or maximise, as appropriate) the chosen cost function.

2.5.1 Deterministic parameter estimation: Nonlinear least squares

A straightforward method for nonlinear model parameter estimation is least squares, whereby the optimal parameter estimates minimise the 'response error' (2.3). This approach requires a numerical algorithm to iteratively supply the parameter estimates such that the cost function can be computed.

2.5.2 Stochastic parameter estimation : Maximum likelihood

Maximum Likelihood is an objective procedure for model parameter estimation that does not suffer from the shortcomings of least squares methods. ML is a statistically rigorous, flexible technique that provides model parameter estimates that are asymptotically unbiased with minimum variance under the assumption that the model residuals are zero mean, Gaussian and uncorrelated. As a result, ML is a popular parameter estimation technique that has been adopted by many scientific disciplines.

Consider the generalised, discrete-time, state-space model representation of a nonlinear system,

$$\boldsymbol{x}(k) = \boldsymbol{f}\{\boldsymbol{x}(k-1), \boldsymbol{u}(k-1), \boldsymbol{\zeta}(k-1)\} \tag{2.43}$$

$$\boldsymbol{y}(k) = \boldsymbol{h}\{\boldsymbol{x}(k), \boldsymbol{e}(k)\} \tag{2.44}$$

where the discrete-time p dimensional vector of observed output variables $\boldsymbol{y}(k) = [y_1(k),\ y_2(k), \cdots,\ y_p(k)]$ are a nonlinear function $\boldsymbol{h}\{\cdot\}$ of state $\boldsymbol{x}(k)$ and measurement noise $\boldsymbol{e}(k)$. The n dimensional state vector, $\mathbf{x}(k) = [x_1(k), x_2(k), \cdots, x_n(k)]$ is a nonlinear function $\boldsymbol{f}\{\cdot\}$ of the state $\boldsymbol{x}(k-1)$, input $\boldsymbol{u}(k-1)$ and system noise $\boldsymbol{\zeta}(k-1)$ vectors at the *k-1*th time instant. Here, the m dimensional input vector $\boldsymbol{u}(k)$ is defined as $\boldsymbol{u}(k) = [u_1(k),\ u_2(k), \cdots,\ u_m(k)]$ and $\boldsymbol{\zeta}(k)$ and $\boldsymbol{e}(k)$ are assumed as independent, normally distributed white noise, with zero mean and covariance $\boldsymbol{Q}$ and $\boldsymbol{R}$ respectively.

Let the matrices $\boldsymbol{Y}(k)$ and $\hat{\boldsymbol{Y}}(k)$ contain the discrete measurements of the system and model output, up to and including time step k, where $k = 0, 1, 2, \cdots, N$.

$$\hat{\mathbf{Y}}(k) = [\hat{\mathbf{y}}(k), \hat{\mathbf{y}}(k-1), ..., \hat{\mathbf{y}}(1), \hat{\mathbf{y}}(0)]^{\mathrm{T}} \tag{2.45}$$

$$\boldsymbol{Y}(k) = [\boldsymbol{y}(k), \boldsymbol{y}(k-1), ..., \boldsymbol{y}(1), \boldsymbol{y}(0)]^{T} \tag{2.46}$$

The joint density function $L(\boldsymbol{\theta}; \boldsymbol{Y}(k))$ can be considered as the probability of obtaining particular values of $\boldsymbol{Y}(k)$, based on the unknown parameters in the vector, $\boldsymbol{\theta}$ which includes the leading diagonals of both $\boldsymbol{Q}$ and $\boldsymbol{R}$. Furthermore, the function $L(\boldsymbol{\theta}; \boldsymbol{Y}(k))$ can be considered as a method of assessing the merits of different estimates of the parameter vector $\hat{\boldsymbol{\theta}}$ given observations in $\boldsymbol{Y}(k)$ and can, therefore, be termed the *likelihood function*,

$$L(\boldsymbol{\theta}; \boldsymbol{Y}(k)) = p(\boldsymbol{Y}(k), \boldsymbol{\theta}) \tag{2.47}$$

Under the assumption that the system $\boldsymbol{\zeta}(k)$ and measurement noise $\boldsymbol{e}(k)$ are normally distributed, the conditional density is also normal. The conditional distribution is, therefore, characterised by the conditional mean, $\hat{\boldsymbol{y}}(k|k-1)$ and variance $\boldsymbol{P}(k|k-1)$; the model output one-step ahead prediction and variance respectively.

$$\hat{\boldsymbol{y}}(k \mid k-1) = E\left[\boldsymbol{y}(k) \mid \boldsymbol{Y}(k-1), \hat{\boldsymbol{\theta}}\right] \tag{2.48}$$

$$\hat{\boldsymbol{P}}(k \mid k-1) = V\left[\boldsymbol{y}(k) \mid \boldsymbol{Y}(k-1), \hat{\boldsymbol{\theta}}\right] \tag{2.49}$$

The one-step ahead prediction error (or innovation) $\varepsilon(k)$ can therefore be defined as

$$\varepsilon(k) = y(k|k) - \hat{y}(k|k-1) \tag{2.50}$$

The likelihood function is defined as,

$$L(\theta; \boldsymbol{Y}(N)) = \prod_{1}^{N} \left[(2\pi)^{-m/2} \det \boldsymbol{P}(k \mid k-1)^{-1/2} \times \exp\left(-\frac{1}{2} \varepsilon(k)^{\mathrm{T}} \boldsymbol{P}(k \mid k-1)^{-1} \varepsilon(k) \right) \right] \tag{2.51}$$

or more conveniently, considered in the following form,

$$\log L(\theta; \boldsymbol{Y}(N)) = -\frac{1}{2} \sum_{1}^{N} \left[\log(\det \boldsymbol{P}(k \mid k-1)) + \varepsilon(k)^{\mathrm{T}} \boldsymbol{P}(k \mid k-1)^{-1} \varepsilon(k) \right] + const \tag{2.52}$$

The Maximum Likelihood estimate is the parameter vector $\hat{\theta}$ that maximises the likelihood function defined in equation (2.51) or minimises the likelihood defined in equation (2.52). The ML function can be computed using the Kalman filter which calculates $\hat{y}(k|k-1)$ and $\boldsymbol{P}(k|k-1)$ based upon estimates of the model parameter vector $\hat{\theta}$, which are passed directly from a numerical optimisation algorithm.

2.5.3 Linearised Kalman Filter

In order to form the *likelihood*, the one-step ahead prediction error and variance are required which can be calculated for discrete-time linear models using the discrete-

time Kalman Filter (KF). Kalman (1960) discussed the recursive solution of the optimal state estimation (or filtering) and prediction problem for linear discrete-time systems. For nonlinear stochastic systems the Linearised Kalman Filter (LKF) can be derived for the optimal state estimation and prediction of nonlinear systems (Jazwinski, 1970).

The recursive prediction of $\hat{y}(k|k-1)$ and $\boldsymbol{P}(k|k-1)$) from model (2.43-2.44) in its current nonlinear form presents an estimation problem of some complexity. To simplify this process, the LKF makes linear approximations of the state and measurement equations by generating a perturbation around their current estimates at the *k*th time instant by calculating the partial derivatives (akin to the Taylor series). The LKF can then be defined in the following prediction-correction form:

Prediction

$$\hat{\boldsymbol{x}}(k+1|k)=\boldsymbol{f}\{\hat{\boldsymbol{x}}(k|k),\boldsymbol{u}(k),\} \tag{2.53}$$

$$\boldsymbol{P}(k+1|k)=\boldsymbol{A}[k+1,k;\hat{\boldsymbol{x}}(k|k)]\boldsymbol{P}(k|k)\boldsymbol{A}^{\mathrm{T}}[k+1,k;\hat{\boldsymbol{x}}(k|k)]+\boldsymbol{Q}(k+1) \tag{2.54}$$

Correction

$$\begin{aligned}\hat{\boldsymbol{x}}(k+1|k+1)=\hat{\boldsymbol{x}}(k+1|k)+\boldsymbol{K}(k+1;\hat{\boldsymbol{x}}(k+1|k))\cdots\\ \times[\boldsymbol{y}(k+1)-\boldsymbol{h}(\hat{\boldsymbol{x}}(k+1|k),k+1)]\end{aligned} \tag{2.55}$$

$$\begin{aligned}\boldsymbol{P}(k+1|k+1)=[\boldsymbol{I}-\boldsymbol{K}\{k+1;\hat{\boldsymbol{x}}(k+1|k)\}\boldsymbol{H}(k+1;\hat{\boldsymbol{x}}(k+1|k))]..\\ \boldsymbol{P}(k+1|k)[\boldsymbol{I}-\boldsymbol{K}\{k+1;\hat{\boldsymbol{x}}(k+1|k)\}\boldsymbol{H}(k+1;\hat{\boldsymbol{x}}(k+1|k))]^{\mathrm{T}}...\\ +\boldsymbol{K}(k+1;\hat{\boldsymbol{x}}(k+1|k))\boldsymbol{R}(k+1)\boldsymbol{K}^{\mathrm{T}}(k+1;\hat{\boldsymbol{x}}(k+1|k))\end{aligned} \tag{2.56}$$

Kalman Gain

$$\begin{aligned}\boldsymbol{K}(k+1;\hat{\boldsymbol{x}}(k+1|k)) = \boldsymbol{P}(k+1|k)\boldsymbol{H}(k+1;\hat{\boldsymbol{x}}(k+1|k))^{\mathrm{T}} \\ \left[\boldsymbol{H}(k+1;\hat{\boldsymbol{x}}(k+1|k))\boldsymbol{P}(k+1|k)\boldsymbol{H}^{\mathrm{T}}\right. \\ \left.(k+1;\hat{\boldsymbol{x}}(k+1|k)) + \boldsymbol{R}(k+1)\right]^{-1}\end{aligned} \tag{2.57}$$

One-step ahead predictions of the measurement equation become:

$$\hat{\boldsymbol{y}}(k+1|k) = \boldsymbol{h}\left(\hat{\boldsymbol{x}}(k+1|k), \boldsymbol{u}(k+1|k), \boldsymbol{\theta}, k+1\right) \tag{2.58}$$

The partial derivatives (or Jacobians) of the state and measurement equations are defined by the matrices $\boldsymbol{A}$ and $\boldsymbol{H}$, where for simplicity, the notation for the kth discrete-time step is not included, even though they are different at each time step.

$$\mathbf{A}_{ij} = \frac{\partial \mathbf{f}_i}{\partial x_j}\left(\hat{\mathbf{x}}(k), \mathbf{u}(k), \boldsymbol{\theta}, k\right) \tag{2.59}$$

$$\boldsymbol{H}_{ij} = \frac{\partial \boldsymbol{h}_i}{\partial x_j}\left(\hat{\boldsymbol{x}}(k), \boldsymbol{u}(k), \boldsymbol{\theta}, k\right) \tag{2.60}$$

The complete stochastic ML parameter estimation procedure is summarised by the schematic given in Figure 2.5.

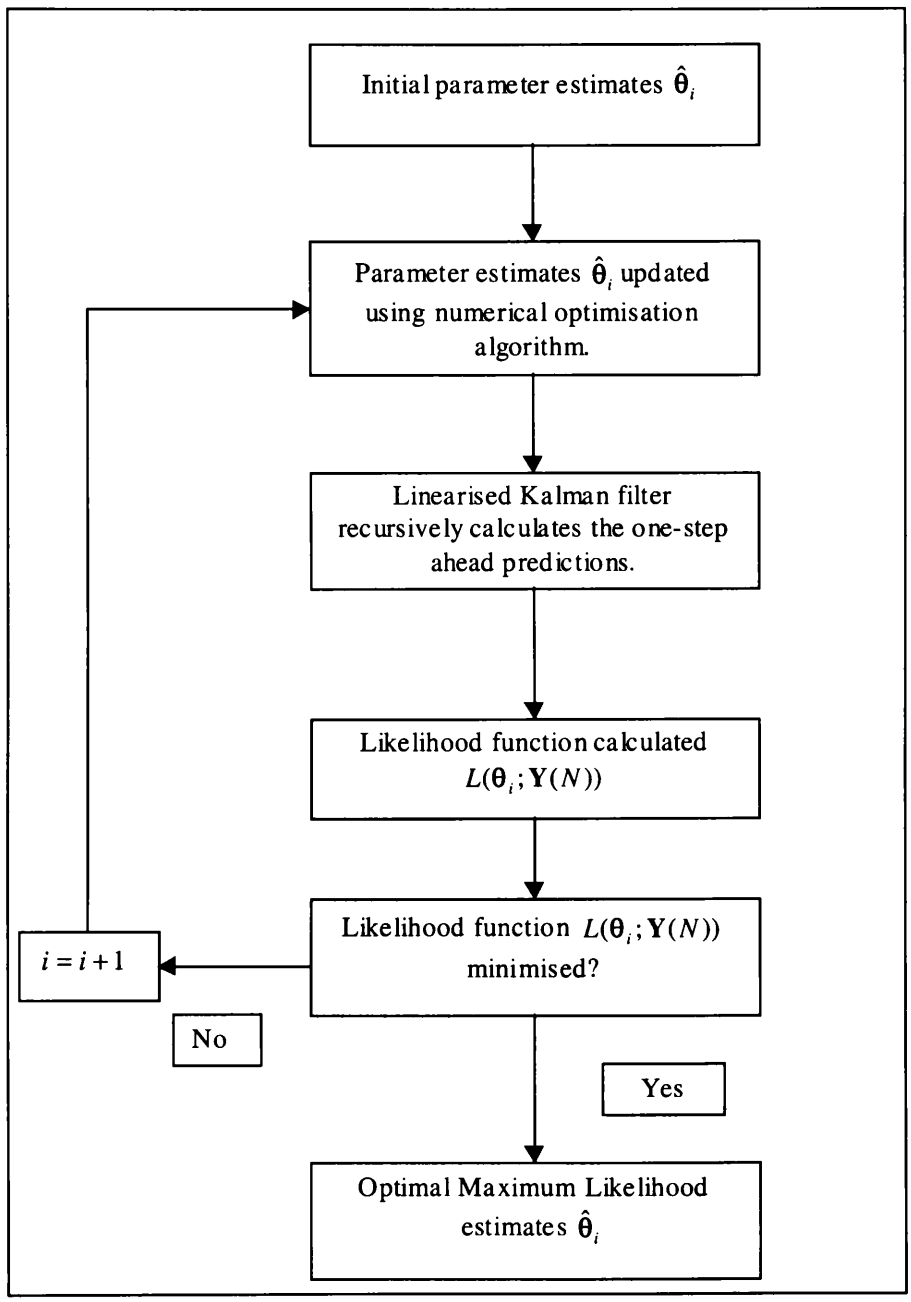

Figure 2.5 Summary of maximum likelihood estimation procedure

2.5.4 ML estimation of continuous-time nonlinear systems

ML parameter estimation of a continuous-time ecosystem model is discussed in Chapter 6. Due to time constraints, a ML parameter estimation software package in

FORTRAN for continuous-time systems (Madsen and Melgaard, 1993) was utilised, rather than writing new software in MATLAB. The Continuous Time Linear Stochastic Modelling (CTLSM) software assumes that the model output observations are taken at discrete-time intervals and, therefore, discretises the continuous-time model state and covariance equations at each iteration for use in the discrete-time LKF. As a result, the calculations involved are computationally very demanding and any subsequent research would require the formulation of the complete ML routines directly in MATLAB, probably by expressing the LKF in continuous-discrete form. In this manner, discretisation would be avoided and the solution would be obtained more efficiently.

2.5.5 Numerical Optimisation Algorithms

Numerical optimisation algorithms provide the parameter estimates that minimise the 'objective functions' of ML and nonlinear least squares. A numerical optimisation algorithm is a logical method that searches the response surface of the parameter hyperspace to locate the best set of parameters that optimises the specified objective function. There are many different numerical optimisation algorithms available (e.g. simulated annealing, genetic algorithms, quasi-Newton algorithms) and each uses a 'strategic' mathematical method to obtain the estimates of the optimal parameters. Most methods work on the principle of creating an initial 'guess' to the optimal solution and then make iterative attempts to enhance this guess following logical procedures defined by the strategy employed. Each strategic method can be crudely categorised as either a local search or global search procedure. Algorithms of both types were utilised for parameter estimation in the studies reported in this thesis. Estimation by nonlinear least squares adopted both a global (genetic algorithm) and a

local (e.g. quasi-Newton based algorithm) search procedure, whilst only a local search method was utilised with ML.

Local search methods utilise the mathematical properties of the response surface (e.g. the gradient) to control the direction and distance of the search from the current location 'downhill' towards the optimal function solution within the parameter hyperspace. Local search methods can be classified as either 'direct search' or 'gradient search' methods. The MATLAB function FMINU is a 'gradient search' type strategy. Gradient methods utilise the difference in cost functional values and the gradient between two points to determine the direction and distance of the future search; and most methods are based on the following equations (Sorooshian and Gupta, 1995),

$$\boldsymbol{\theta}_{l+1} = \boldsymbol{\theta}_l - \rho \cdot \mathbf{A} \cdot \nabla \boldsymbol{\theta}_l \tag{2.61}$$

where $\boldsymbol{\theta}_{l+1}$ is the generated new point in the parameter hyperspace; $\boldsymbol{\theta}_l$ is the current point; $\nabla \boldsymbol{\theta}_l$ is the gradient matrix at the present point; ρ determines the step size and $\mathbf{A}$ is a specifically chosen square matrix. It can be shown mathematically, that if $\mathbf{A}$ is a positive definite matrix, then the vector from $\boldsymbol{\theta}_l$ to $\boldsymbol{\theta}_{l+1}$ will give a better cost functional value (Sorooshian and Gupta, 1995). In the case of quasi-Newton methods, $\mathbf{A}$ is defined using an approximation of the Hessian matrix. If the newly defined point in the parameter space provides a lower functional value than the previous point, then it replaces the previous point. If the new functional value is higher, the step size is reduced and a new point is evaluated. At the optimal point in the parameter hyperspace the gradient will be close to zero and the search is terminated. Local

search methods are limited in cases where the parameter hyperspace has a complex and irregular 'response surface' such that they often terminate in local rather than global minima. For this reason, the success of local search methods is often dependent upon the location of the search initiation.

Global search methods avoid this sub-optimal, local optimisation by continually looking away from the local optima that the algorithm is working towards. A genetic algorithm (GA) (see e.g. Goldberg, 1989) is a stochastic global search method which uses the biological concepts of 'natural evolution' and the 'survival of the fittest' to obtain optimal parameter estimates. Genetic algorithm's work with a collection of randomly selected populations of potential solutions rather than a single solution. Each population is comprised of a number of chromosomes containing genes, which represent each parameter to be estimated. Each individual is assessed within a defined objective function and assigned a 'fitness value' which in turn, determines whether it is selected for breeding. The fittest individuals are selected and a undergo a breeding process using crossover and mutation operators to generate a new population of chromosomes which are subsequently evaluated. The process of mutation randomly alters the configuration of the chromosome maintaining its genetic diversity in order to prevent premature convergence as a consequence of a local optimum. The GA evolves a population of chromosomes over many cycles or generations until a specified termination criterion is met.

2.6 CONCLUSION

This chapter has introduced the DBM identification and estimation techniques that have been adopted in the research presented in this thesis to model linear and nonlinear systems. The parameters of *linear* TF's are generated using the Simplified Refined Instrumental Variable (SRIV) algorithm which provides consistent, asymptotically unbiased estimates through adaptive pre-filtering of the IV vector. The best models are subsequently chosen using the R_T^2 and *YIC* identification criteria. The DBM approach to modelling nonlinear systems utilises objective non-parametric time varying parameter (TVP) estimation, based on the recursive Fixed Interval Smoothing (FIS) algorithm, to detect and identify any significant nonlinearities in the time series data. The identified nonlinearity is efficiently parameterised, so defining the structure of the nonlinear model, using a state dependent parameter modelling (SDPM) procedure. Final estimates of the nonlinear model parameters are generated using Maximum Likelihood, or least squares approaches.

Chapters 3, 4 and 5 utilise the DBM approach to identify and estimate stochastic nonlinear TF models of ecological and hydrological systems, whilst in Chapter 6, the parameters of a deterministic, continuous-time, ecological model are estimated, within a data assimilation framework, using least squares and Maximum Likelihood methods.

CHAPTER 3

DBM MODELLING OF *LUCILIA CUPRINA* POPULATIONS

In Chapter 2, the data-based mechanistic (DBM) approach to modelling nonlinear systems has been introduced. This chapter demonstrates the efficacy of this approach by utilising these procedures in an ecological context, to model population data of the Australian sheep blowfly, *Lucilia cuprina*, that were collected by A.J. Nicholson.

3.1 NICHOLSON'S LUCILIA CUPRINA EXPERIMENTS

Throughout the 1950's the Australian entomologist A. J. Nicholson performed a comprehensive series of laboratory experiments to investigate the single-species population dynamics of the *Lucilia cuprina*, under various pre-determined conditions (Nicholson, 1950; 1954; 1957). The essential feature of these experiments was that the blowfly population were allowed to develop freely, regulated only by the rate of food supply (constant or fluctuating), to either the adults or larvae. As a result, the population experienced different forms of competition, e.g. for food or egg laying space. Nicholson's acclaimed experiments have caused a wide interest in population

ecology and his data have appeared in numerous ecology texts as examples of oscillatory populations.

In what has probably become the most analysed of Nicholson's experiments, a population of adult blowfly were fed a constant but limited daily supply of ground liver (0.5 g) whilst the larvae, isolated from the adult blowflies, had access to an unlimited supply of larval food. Measurements of the number of adult blowflies and the eggs laid per day were taken by Nicholson every two days for approximately one year. The oscillatory nature of the population dynamics are illustrated in Figure 3.1, exhibiting large changes in amplitude but maintaining an approximately constant period.

Nicholson (1954) attributes these oscillations to population density-related competition for food. The highest egg generation rate occurs when the adult population is very low. As the population increases so does the competition for food until it becomes sufficiently severe that all or some adults no longer lay eggs. Where natural mortality has caused the population to decrease to a level when competition for food is not as intense, some individuals can secure sufficient food to enable egg generation. After approximately two weeks, the eggs develop into mature adults and the population density rises and the cycle repeats itself once again. Nicholson also concludes that if the transformation of eggs to adults was instantaneous, the system would be non-oscillatory and the population would increase until an *equilibrium density* was reached.

Unfortunately, Nicholson's original data have been lost by his trustee, Professor Don McNeil and consequently the data used in this research and shown in Figure 3.1 were obtained from careful digitisation of a figure taken from Nicholson's original manuscript (Figure 3, Nicholson, 1954).

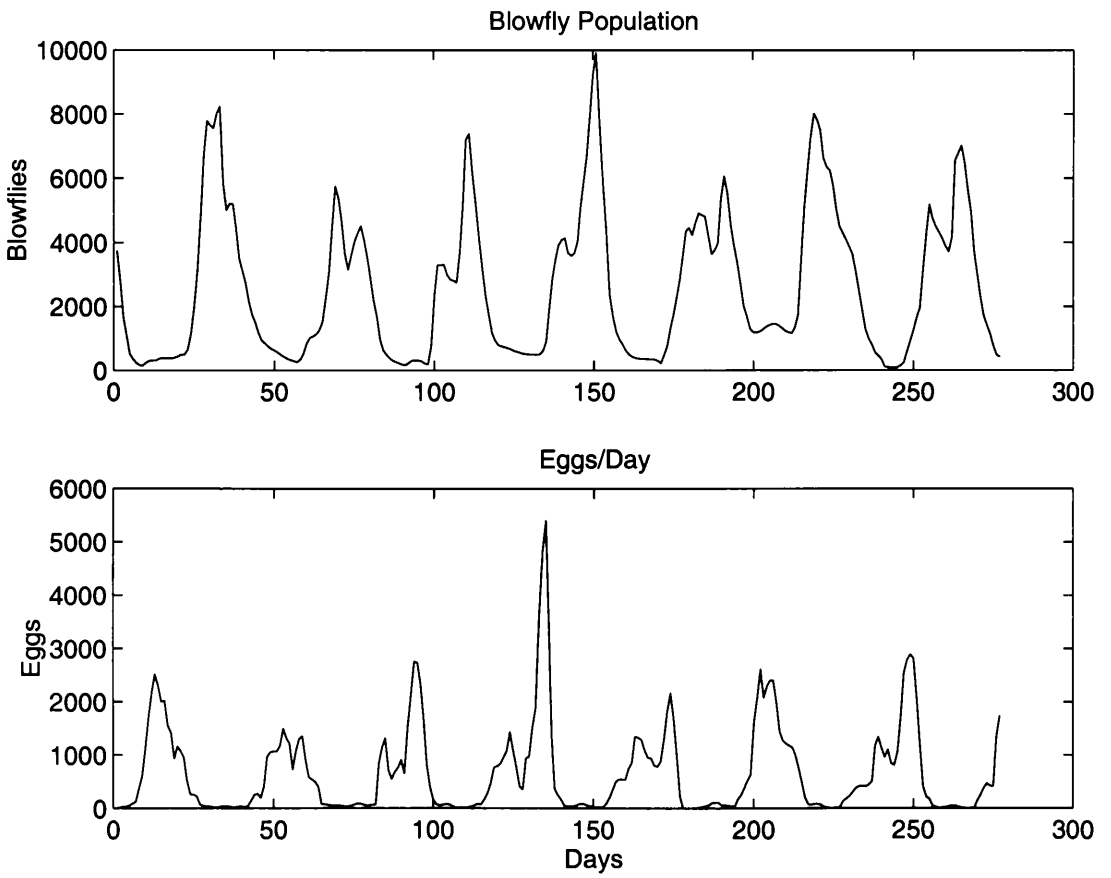

Figure 3.1 Numbers of blowfly *Lucilia cuprina* in a population cage.

Adult blowfly (Top) and eggs laid per day (Bottom). Data sampling interval is 1 day.

3.2 MODELLING NICHOLSON'S BLOWFLY

Since the publication of these data, numerous mathematicians and ecologists have analysed the time series concentrating mainly on capturing the *dynamics of the blowflies.* This section briefly summaries the results of these earlier studies and comments upon their relative success.

May (1973) reported that the following, simple, time-delayed logistic growth equation was capable of generating the cyclic or quasi cyclic population fluctuations observed in the blowfly time series,

$$\frac{dN(t)}{dt} = rN(t)\left[1 - \frac{N(t-T_D)}{K}\right] \tag{3.1}$$

where the blowfly population $N(t)$ has a maximum growth rate r, and maximum sustainable population K, regulated by the time delay T_D. Although an approximate fit to the adult population time series can be achieved using this model, its acceptability as an adequate representation of the system dynamics is undermined by two important factors.

The most severe objection to the model is due to the 9 day time delay T_D estimated by May (1973), which he argued approximated to the delay observed in the time series. However, careful examination of the data show the real time delay to be in the region of 15 days; a significant discrepancy. Secondly, the delayed logistic growth equation is unable to reproduce the two periods of adult reproductive activity observed in each adult population cycle, characterised by the double peaks in the data. As a result, this

model gives an inadequate representation of the food limited adult population dynamics.

Other attempts to model the blowfly data over the last 30 years include studies by Maynard Smith (1974), Poole (1978), Readshaw and Cuff (1980) and Brillinger *et al.*, (1980). Although various advanced and relatively complex methodologies have been employed by these authors, no overall satisfactory fits to the time series data, or adequate explanations of the mechanisms underlying the blowfly population dynamics have been achieved.

The most recent research by Gurney *et al.*, (1980; 1983) and Nisbet and Gurney (1982) has re-addressed the modelling issue. Recognising the limitations of earlier models, (e.g. May, 1973), Gurney *et al.*, (1980) formulated a new deterministic model, which encompassed the *complete egg-blowfly system* by incorporating a component reflecting the reproductive activity of the blowflies, based on mass balance principles. The adult population $N(t)$ is determined by adult death $D(t)$ and recruitment $R(t)$ rate which, in turn, are both functions of the total adult population at any one time. Therefore, the adult population dynamics can be expressed in continuous-time, by the following differential equation.

$$\frac{dN(t)}{dt} = R(t) - D(t) = R(t) - \delta N(t) \tag{3.2}$$

The recruitment rate function $R(t)$ can incorporate age dependent terms to account for the different stages of egg-mature blowfly growth, but this added complexity can be avoided by following three basic assumptions:

1. The rate of egg production is dependent on population size.
2. T_D is the exact time taken for eggs to develop into mature adults.
3. The likelihood of an egg developing into a mature adult depends upon the number of competitors of the same size.

These assumptions lead to the redefinition of equation (3.2) to this form,

$$\frac{dN(t)}{dt} = R\big(N(t-T_D)\big) - \delta N(t) \tag{3.3}$$

Whilst the death rate function $N(t)$ is controlled by a simple time and density independent value δ, the parameterisation of the recruitment rate function $R(t)$ is far more sophisticated. It was formulated with the aim of explaining the Nicholson blowfly data well and providing a model for generic use with a wider class of insect population models. The authors deduced, *heuristically*, the form of the function from the visual examination of the blowfly data, which is based on the following assumptions (Gurney *et al.*, 1980).

1. The adult recruitment rate is directly proportion to the egg laying rate at time $(t-T_D)$.
2. Where food is limited to the adult blowflies, the total egg laying rate will be affected by the population density and adults will begin to compete for food. It is likely that at high population densities, the food intake per blowfly will be reduced to a level that is just sufficient to maintain their existence only, completely eliminating any possible egg production. These nonlinear effects can be clearly

observed in the time series data, as shown in Figure 3.1, where egg production drops rapidly to zero when high adult populations exist and increase when the population is small.

3. In light of this, a suitable recruitment rate function $R(N)$, must decline to zero at both low and high populations, but rise to a single maximum at an intermediate population size. Gurney *et al.,* (1983) suggested that $R(N)$ could be simply represented by an exponential relationship, defining the complete dynamic model as,

$$R(N) = PN\exp(-N / N_o) \tag{3.4}$$

$$\frac{dN(t)}{dt} = PN(t-T_D)\exp\{-N(t-T_D)/N_D\} - \delta N(t) \tag{3.5}$$

where N_o is the population size that produces the maximum reproduction rate and P is the maximum per capita egg generation rate.

This final deterministic model provides a good fit to the time series data, producing self-sustaining oscillations, and it has since been extended to incorporate age structure within the blowfly population (Nisbet and Gurney, 1982). Through quantitative tests, Gurney *et al.,* (1980) were also able to identify that the population fluctuations were self-sustaining limit cycles rather than driven quasi-cycles.

3.3 DBM MODELLING OF NICHOLSON'S BLOWFLY

The preceding section has reviewed attempts to model Nicholson's blowfly time series by adopting a deterministic approach. With the exception of Gurney's heuristic model, the remaining models have failed to effectively characterise the insect population dynamics.

In this section, a novel analysis of the data is presented which utilises the statistical DBM modelling methodology, to objectively investigate the form of the *complete egg-blowfly relationship* directly from the time series. In particular, emphasis is placed on exploiting the DBM approach to identify the form of the feedback nonlinearity (blowfly-eggs) present in the data and to evaluate whether the resulting relationship has similarities to the function *heuristically* identified by Gurney *et al.,* (1980). In fact, this paper was only consulted after the DBM analysis had successfully identified the state dependent form of the feedback nonlinearity.

The DBM analysis involves three main stages: firstly, a linear transfer function (TF) model of the forward path is identified between eggs laid per day and adult blowflies. Secondly, *time varying parameter* (TVP) and *state dependent modelling* (SDM) techniques (see Chapter 2) are utilised to identify the nonlinear feedback path. This separation of the forward and feedback path identification is possible, in this case, by the large time delay in the forward path which effectively isolates the characteristics in each path for the purpose of identification. However, at the third stage of the analysis, the parameters of the full model are optimised by taking both a deterministic

approach, using least squares; and a stochastic approach based on Maximum Likelihood (ML).

3.3.1 Forward path linear TF model

The first stage of the analysis is to objectively identify a discrete-time linear TF model between eggs laid per day $u(k)$ and the adult blowflies $y(k)$. The best 6 linear TF models, identified and estimated using the Simplified Refined Instrumental Variable (SRIV) algorithm, are shown in Table 3.1 with their corresponding statistical fitting criterion.

Table 3.1 Best identified linear TF models listed in order of R_T^2
The sampling interval of the data is one day

Den	Num	Delay	*YIC*	R_T^2	*AIC*
1	1	15	-8.722	0.877	13.376
1	1	14	-8.693	0.869	13.441
2	1	12	-7.266	0.858	13.526
2	1	11	-7.787	0.840	13.647
1	1	16	-7.954	0.838	13.651
1	1	13	-8.178	0.832	13.692

The model giving the best fit to the data ($R_T^2 = 0.877$), with the most negative YIC (and lowest AIC) is first order with a pure time delay of 15 days and is defined by the following TF in the backward shift operator,

$$y(k) = \frac{b_0}{1 + a_1 z^{-1}} u(k-15) + \zeta(k) \tag{3.6}$$

where $\zeta(k)$ represents the complete unmodelled component. The SRIV algorithm estimates the parameters as $\hat{a}_1 = -0.7346(0.0097) \quad \hat{b}_0 = 0.9706(0.033)$, where the standard errors are given in parentheses. The $R_T^2 = 0.877$ shows that 87.7% of the blowfly time series is explained by the model,

$$\hat{y}(k) = \frac{0.9706}{1-0.7346z^{-1}}u(k-15) \tag{3.7}$$

where it will be noted that the model output $\hat{y}(k)$ is computed from $u(k)$ without any reference to the $y(k)$ measurements. The coefficient of determination based upon the one-step-ahead predictions, i.e.

$$\hat{y}_p(k) = -0.7346y(k-1) + 0.9706u(k-15) \tag{3.8}$$

is much closer to unity at $R^2 = 0.949$. Given these high coefficients of determination, it is not surprising that the output of the simulation model (3.7) fits the data very well, as shown in Figure 3.2. Of course, the output of the prediction model (3.8) fits the data very well indeed.

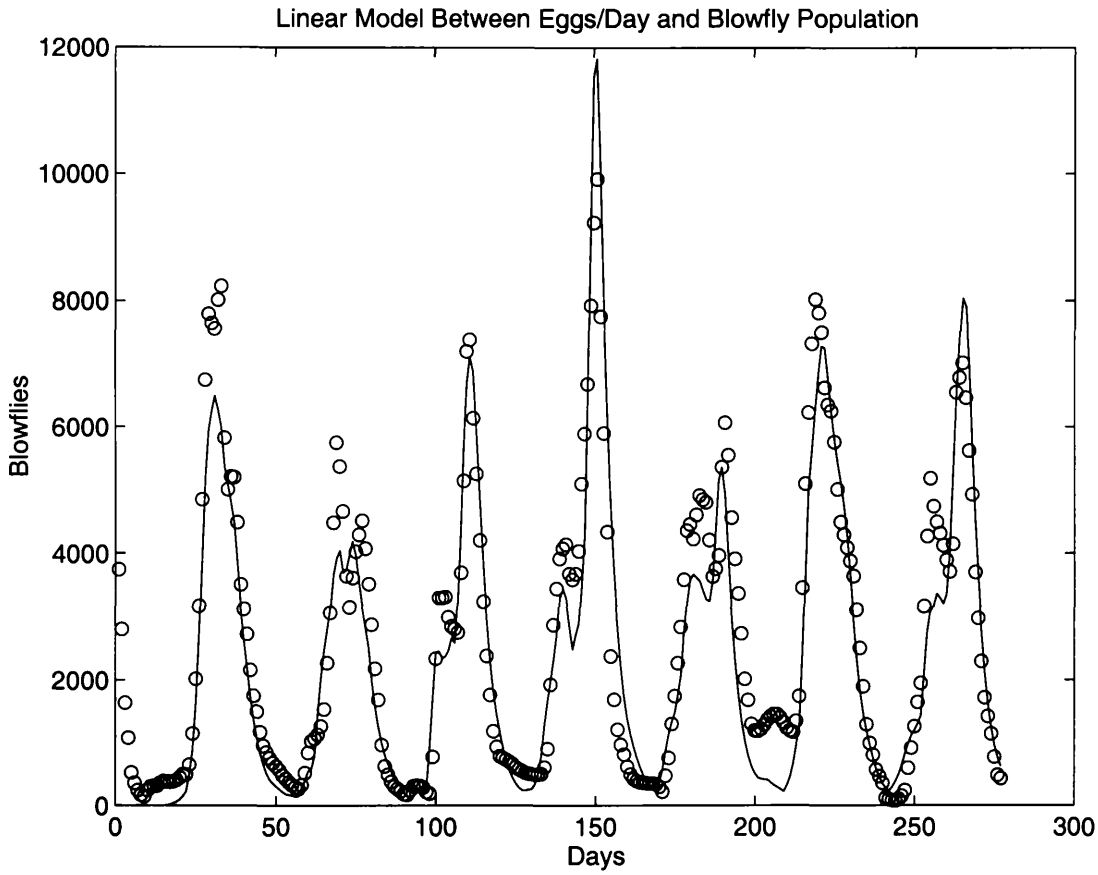

Figure 3.2 Linear TF model between eggs/day and blowflies:

Data (circles); Model (full line).

The *initial* parameter estimates of this linear model have a clear ecological/mechanistic interpretation. The denominator parameter a_1 and numerator parameter b_0 characterise the daily survival rate of the adult blowflies and eggs respectively, indicating that on average 73% of the adult blowflies survive to the next day whilst 97% of the eggs survive to the 15^{th} day. Freeman (1981) also adopted a TF approach to model the forward path of these data and achieved comparable results.

The behaviour of this model can be evaluated more thoroughly by deterministic simulation; considering the model response to an impulse input of eggs. Figure 3.3 presents the output of the model to an impulse of 100 eggs at $k = 0$, where the initial adult blowfly population at this time instant is 100. As a result of the 15 day development period required for the eggs to mature into adults, the population fails to grow and decreases exponentially due to natural mortalities. However, after 15 days there is an instantaneous addition of new blowflies to the population, once the egg/larvae maturation period has been exceeded. Thereafter, the population declines again in an exponential manner.

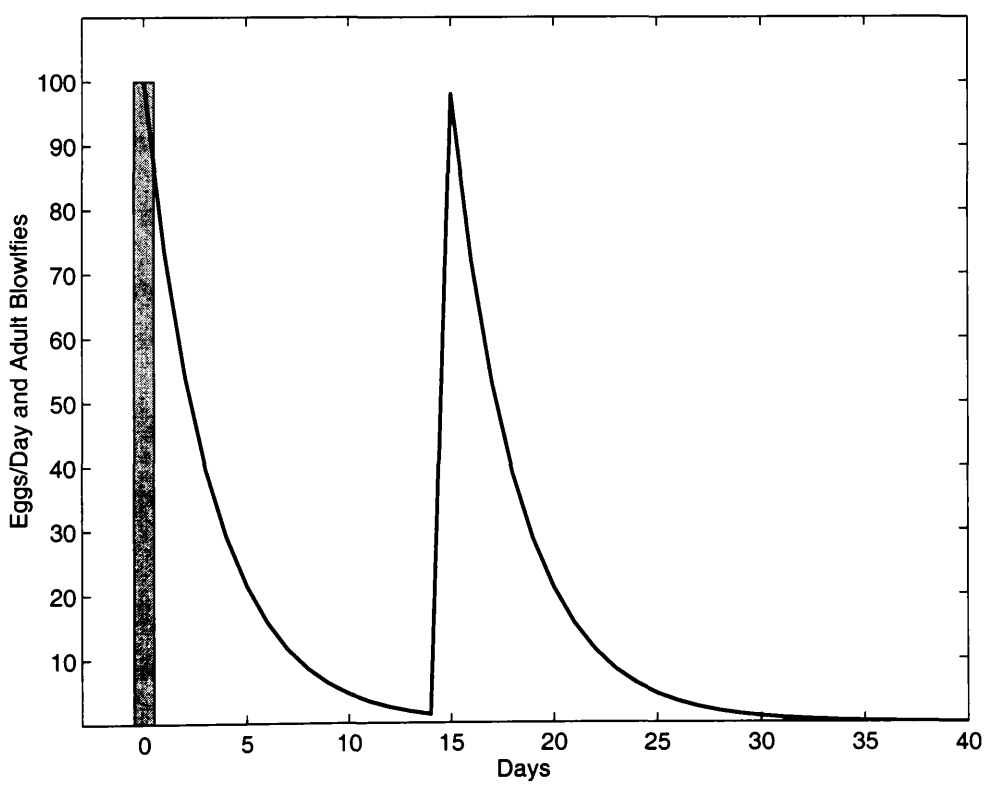

Figure 3.3 Model response to an impulse of 100 eggs (block) assuming an initial blowfly population of 100 adults (line).

Having identified a TF model that successfully reproduces the behaviour of the blowfly population in the forward path, the next stage of the analysis investigates the feedback relationship between blowflies and eggs. An initial investigation, assuming a linear feedback relationship, is unsuccessful in characterising the system dynamics and the blowfly population simply dies out. This behaviour is not surprising, as visual examination of the data and research by Gurney identifies the process to be nonlinear and, therefore, it requires an alternative formulation.

3.3.2 TVP/SDPM Identification of the feedback nonlinearity

The data clearly shows some form of nonlinear inverse relationship existing between the egg and blowfly series; maximum eggs are laid when the blowfly population is at its lowest. An objective, novel and interesting approach to investigate the form of this nonlinearity, is to adopt the *time varying parameter* (TVP) and *state dependent parameter modelling* (SDPM) methodology as outlined in the previous chapter.

The feedback relationship between the blowflies and eggs can be represented by the following TVP equation:

$$u(k) = a(k)y(k) \tag{3.9}$$

A non-parametric estimate $\hat{a}(k/N)$ of $a(k)$ and the associated covariance matrix $\hat{P}(k/N)$ can be obtained from the *Fixed Interval Smoothing* (FIS) algorithm under the assumption of a random walk (RW) variation, with a *noise variance ratio* (NVR) numerically optimised by ML. State dependent analysis of the estimated TVP,

indicates that an inverse cubic relationship exists between $\hat{a}(k/N)$ and $y(k)$ although this relationship is not particularly well defined. However, a much better defined state dependent relationship can be identified by sorting $y(k)$ in ascending numerical order, reordering $u(k)$ in an identical manner, and re-estimating $\hat{a}(k/N)$. The new non-parametric estimate of $\hat{a}(\circ/N)$ and standard error bounds obtained using the optimised FIS algorithm is shown in Figure 3.4.

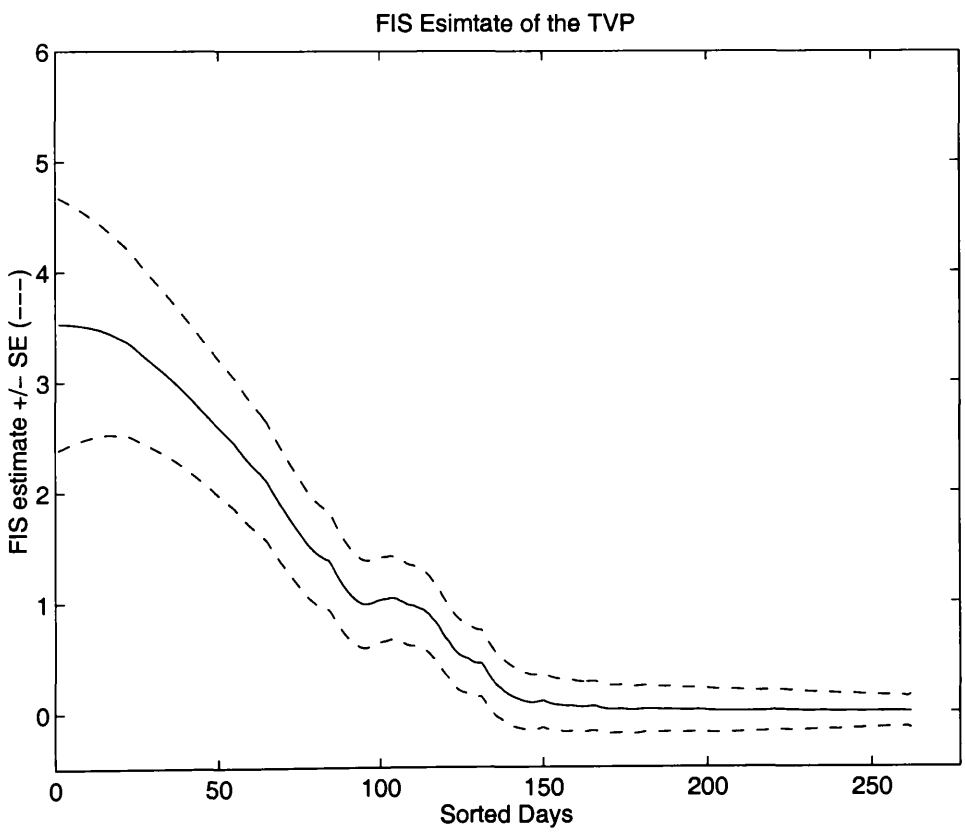

Figure 3.4 FIS estimate $\hat{a}(\circ/N)$ of TVP from sorted data (full line) and standard error bands (dashed line).

The refined and improved state dependent relationship is identified by plotting $\hat{a}(\circ / N)y(k)$, the non-parametric estimate of $u(k)$, based on $y(k)$, versus $y(k)$ and this is illustrated in Figure 3.5.

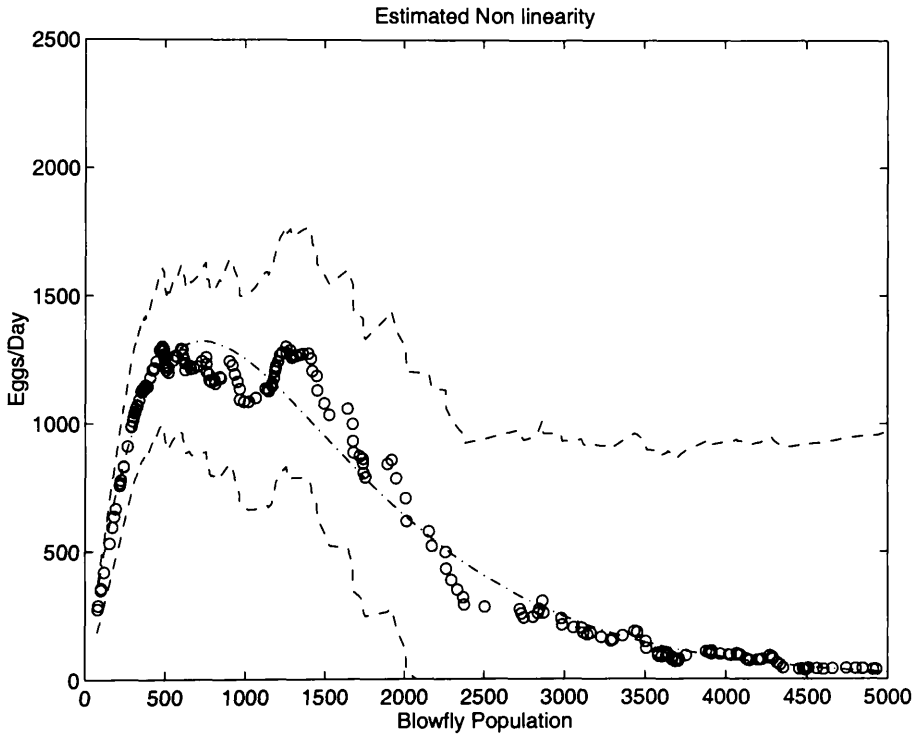

Figure 3.5 Estimated feedback nonlinearities: FIS non-parametric estimate of $u(k)$: $\hat{a}(\circ / N)y(k)$ versus $y(k)$ (circles) with standard error bands (dashed); parameterised WLS estimate of non-parametric result (dot-dashed).

The relationship suggests that at low blowfly populations, the egg production rate is approximately proportional to the population. As the population increase, egg production decreases exponentially away from its maximum, until at high populations, the daily egg production rate almost drops to zero. It is significant that the form of the feedback nonlinearity, objectively and effectively identified from the data by following

the DBM procedures, matches and confirms the function deduced by Gurney who adopted the more conventional 'hypothetico deductive' approach. There are different conceivable ways of parameterising the nonlinearity, but one possibility is clearly the parameterisation suggested by Gurney *et al.*, (1980), which is presented below using alternative nomenclature,

$$\hat{a}(k/N) = g\exp\left(-\frac{1}{fb(k).M}\right) \tag{3.10}$$

Here the food *per capita* $fb(k)$, is defined as $fb(k) = y(k)/0.5$, assuming equal division of the constant daily liver supply of 0.5 g; and M is the population size that has the optimal reproductive rate. Under this assumption, the feedback nonlinearity can be defined as,

$$u(k) = \hat{a}(k/N).y(k) \tag{3.11}$$

$$u(k) = g.y(k).\exp\left(-\frac{y(k)}{0.5M}\right) \tag{3.12}$$

Initial estimates of the parameters in equation (3.12) can be obtained utilising WLS (see Chapter 2). A comparison of the resulting parametric model fit and the non-parametric estimate of the nonlinearity is shown in Figure 3.5, where the estimated parameters are; $g = 4.91$ and $M = 1465$. This confirms the validity of the suggested state dependent function, although other parameterisations that accommodate the non-parametric estimate of the nonlinearity are possible (see e.g. Young, 1999).

Having *identified* the *structure* of the feedback nonlinearity, the parameters of the egg-blowfly closed loop nonlinear TF model, defined by equations (3.13-3.14), can be estimated. Notice that, with the addition of the feedback component of the model, the parameter b_0 is unidentifiable within the whole model structure since it is effectively replaced by g, and is eliminated from the equation (3.14). The structure of the continuous-time version of this discrete-time model is presented in equation (3.15) and compares exactly with Gurney's heuristically derived model (see equation 3.5).

$$u(k) = g.y(k).\exp\left(-\frac{y(k)}{0.5M}\right) \tag{3.13}$$

$$y(k) = a_1 y(k-1) + u(k-15) \tag{3.14}$$

$$\frac{dy(t)}{dt} = ay(t) + gy(t-15)\exp\left(\frac{-y(t-15)}{0.5M}\right) \tag{3.15}$$

The DBM concept to modelling has *objectively* identified a blowfly model from the time series data which, in this instance, has acted as an independent validation of the model suggested by Gurney.

At this stage in the DBM analysis, deterministic and stochastic methods of optimisation can be used to estimate the parameters in the complete TF model (equations 3.13-3.14) and the application of these procedures will be discussed in the next two sections.

3.4 OPTIMISATION OF THE DETERMINISTIC BLOWFLY MODEL

A deterministic method of estimating the parameters in equations (3.13-3.14) can be carried out by using the Least Squares (LS) criterion shown in equation (3.16); where the sum of the combined squared blowfly and egg residuals are minimised,

$$J = \sum_{k=1}^{k=N} \left[e_y(k)^2 + e_u(k)^2 \right] \tag{3.16}$$

where, $e_y(k) = y(k) - \hat{y}(k)$ and $e_u(k) = u(k) - \hat{u}(k)$.

Other related LS criteria could be used with different weighting on the $e_y(k)^2$ and $e_u(k)^2$ terms in (3.16) but, since this deterministic approach is rejected later, it will suffice to consider this simple LS criterion (3.16). Numerical optimisation of this function using the MALTAB (Mathsworks, 1992) nonlinear least squares algorithm provides the following parameter estimates and associated standard errors: $a_1 = 0.8177(0.002)$, $g = 8.937(0.443)$ and $M = 824(30.45)$.

As shown in Figure 3.6, the optimised model provides a good fit to the blowfly and egg time series, with coefficient of determination values of $R_T^2 = 0.774$ and $R_T^2 = 0.559$ respectively.

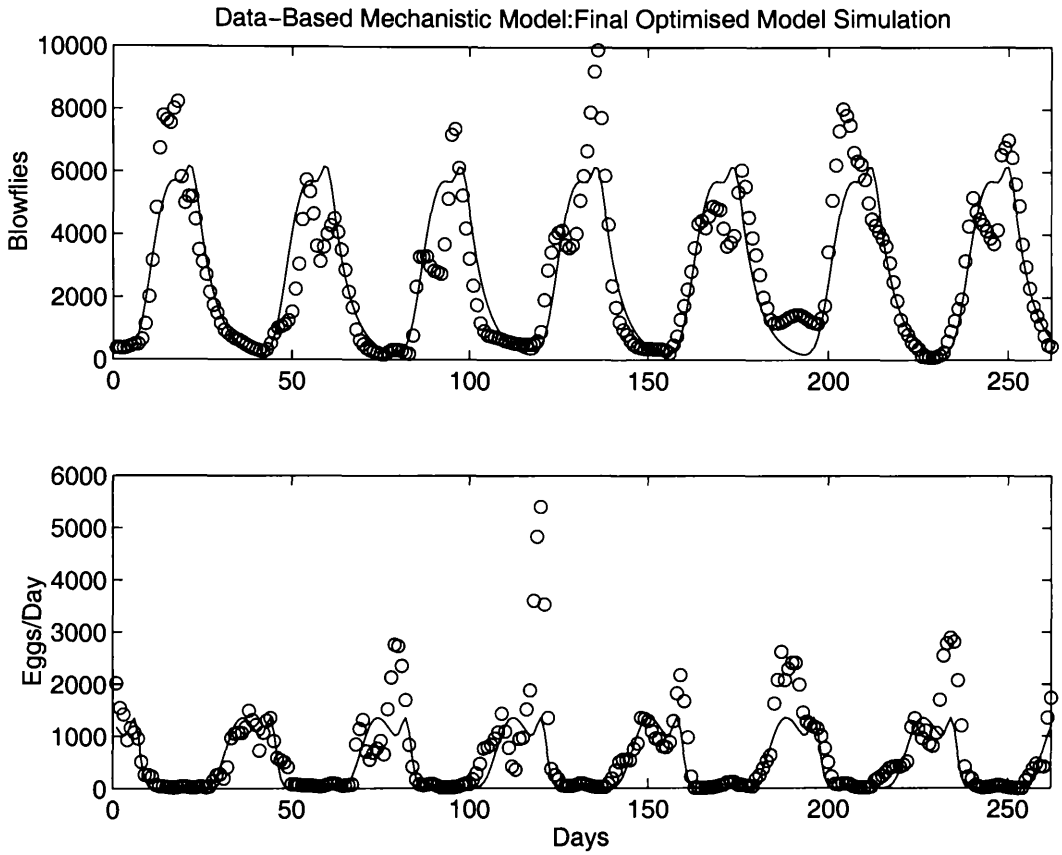

Figure 3.6 The optimised deterministic nonlinear model output (full line) compared with the noisy measured data (circles).

The efficacy of the estimated model can be evaluated further by comparing the feedback nonlinearity predicted from the deterministic optimised model, directly with the non-parametric FIS estimate, as shown in Figure 3.7. At low blowfly populations, the deterministic optimised nonlinearity gives a reasonable match to the FIS estimated number of eggs laid per day. However, as the population size increases, the number of eggs laid per day predicted by the deterministic nonlinearity, drops well below the FIS non-parametric standard error bounds. This discrepancy throws some doubt on the deterministic results and this is discussed in subsequent sections of this chapter.

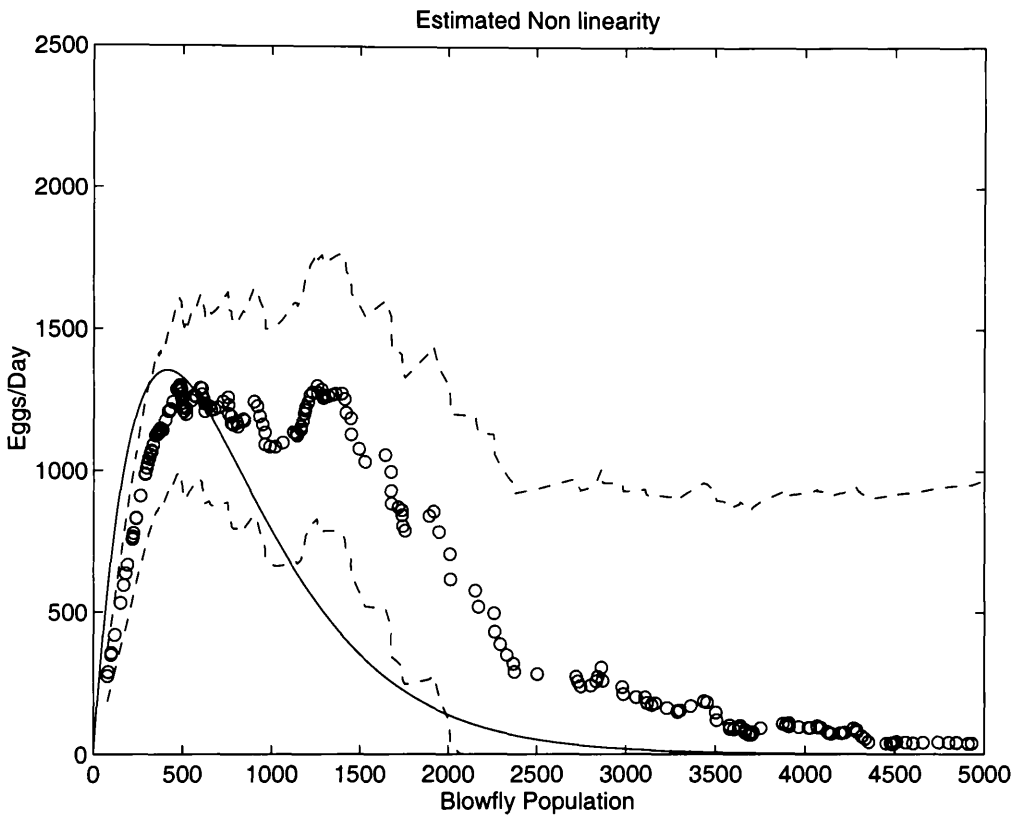

Figure 3.7. Estimated feedback nonlinearities: FIS estimate (circles) with standard error bands (dot dashed) and the deterministic, optimised model output (full).

The residuals of the optimised model offer additional information to assess model efficacy. If the errors of the model *are not* a sequence of serially, uncorrelated, random white noise, this may indicate that either the estimated parameters are deficient, which may be directly attributable to the estimation process, or that the model structure is not a good representation of the system. The model errors can be analysed by using the autocorrelation (AC) and partial autocorrelation (PAC) functions (see e.g. Box and Jenkins, 1976) to evaluate whether the series is either a serially uncorrelated sequence of white noise, or a sequence of coloured noise with definite structure.

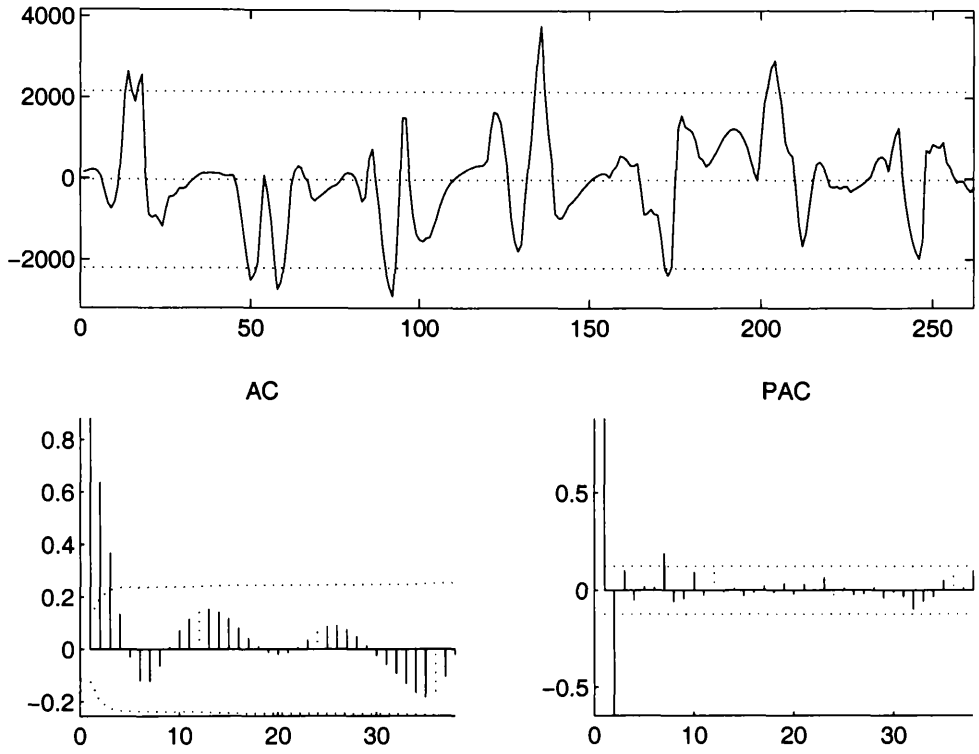

Figure 3.8 Model errors (top). The autocorrelation and partial autocorrelation functions for the least squares optimised model are shown in the bottom left and bottom right figures respectively. An instance above the dashed line in these figures, indicates the model errors are correlated.

Figure 3.8 presents the AC and PAC functions of the model residuals and reveals that the series is, indeed, autocorrelated with a 'coloured' temporal pattern. This emphasises that the model is limited in its present state and indicates that alternative parameterisation or model structures should be investigated.

3.5 STOCHASTIC OPTIMISATION OF THE DETERMINISTIC BLOWFLY MODEL

A more statistically rigorous method of parameter estimation, can be achieved by stochastic optimisation based on Maximum Likelihood (ML) theory (see Chapter 2). By converting the TF model into stochastic state space form (equations 3.17-3.19), the model parameters can be estimated by numerical optimisation, minimising the Gaussian conditional density function (likelihood) generated by prediction error decomposition.

For nonlinear models, the conventional procedure for generating the one-step ahead predictions required for ML estimation, would be to utilise the *Linearised Kalman Filter* (LKF). As outlined in Chapter 2, a software package is commercially available for estimating parameters of continuous-time nonlinear state space models (CTLSM), including time varying parameters, based on ML. As the discrete-time nonlinear TF (equations 3.13-3.14) can be easily converted into continuous-time, (see e.g. Young, 1993) the CTLSM software would, therefore, appear to be a convenient method for estimating the parameters of the blowfly model. Unfortunately, because the CTLSM software does not have the facility to deal with pure time delays in the state, it is an unsuitable method for estimating the parameters of the blowfly model.

To overcome this problem, a *novel* parameter estimation procedure based on ML has been developed utilising the ordinary discrete-time linear *Kalman Filter* with state dependent parameters. At *each recursive time step,* the nonlinear component of the stochastic state space model $j(k)$ (equation 3.19 below) is computed *before* state

prediction and correction is undertaken using the current estimates of the unknown parameters. In this manner, the nonlinear model is treated as a linear, time varying parameter model which is able to exploit the fact that the Kalman filter is formulated in such a time variable form. Of course, this time variation in the parameters is in fact caused by a state dependency but this causes no problem in this case.

The 15th order stochastic state space model used in this manner takes the form,

$$\begin{aligned} \boldsymbol{x}(k) &= \boldsymbol{F}\boldsymbol{x}(k-1) + \xi(k) \\ y(k) &= \boldsymbol{H}\boldsymbol{x}(k) + e(k) \end{aligned} \tag{3.17}$$

where the state vector is defined as,

$$\boldsymbol{x}(k) = [y(k), y(k-1), y(k-2), y(k-3), ..., y(k-13), y(k-14)]^T \tag{3.18}$$

the time varying parameter is given by,

$$j(k) = g.\exp\left(-\frac{y(k-15)}{0.5M}\right) \tag{3.19}$$

the 15×15 state transition matrix $\boldsymbol{F}$ and the observation matrix $\boldsymbol{H}$ are,

$$\boldsymbol{F} = \begin{bmatrix} -a & 0 & 0 & \cdots & 0 & 0 & j(k) \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

$$\boldsymbol{H} = [1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0]^T$$

and the system $\xi(k)$ and measurement noise $e(k)$ processes are characterised by the *noise variance ratio* (*NVR*) matrix $Q_1 = \dfrac{\sigma_\xi^2}{\sigma_e^2}$.

A quasi-Newton optimisation algorithm (MATLAB's FMINU function) was utilised to generate the estimates of the four parameters to be optimised (a_1, g, M and Q_1), needed to minimise the Gaussian likelihood function (see equation 2.52). Figure 3.9 presents the one-step ahead predictions that are obtained from the optimised model, shown in equations (3.20-3.22), and the resulting limit cycle. The one-step ahead predictions closely match the data and the limit cycle characterises both the amplitude and period of the data, with each adult population cycle showing two periods of reproduction. Note that the optimised parameter values are not very different to those obtained from the deterministic optimisation in Section 3.4.

$$u(k) = 9.958x(k).\exp\left[-\frac{x(k)}{0.5\times 849.163}\right] \tag{3.20}$$

$$x(k) = 0.8203x(k-1) + u(k-15) + \xi(k) \tag{3.21}$$

$$y(k) = x(k) + e(k) \quad Q_1 = 0.0049 \tag{3.22}$$

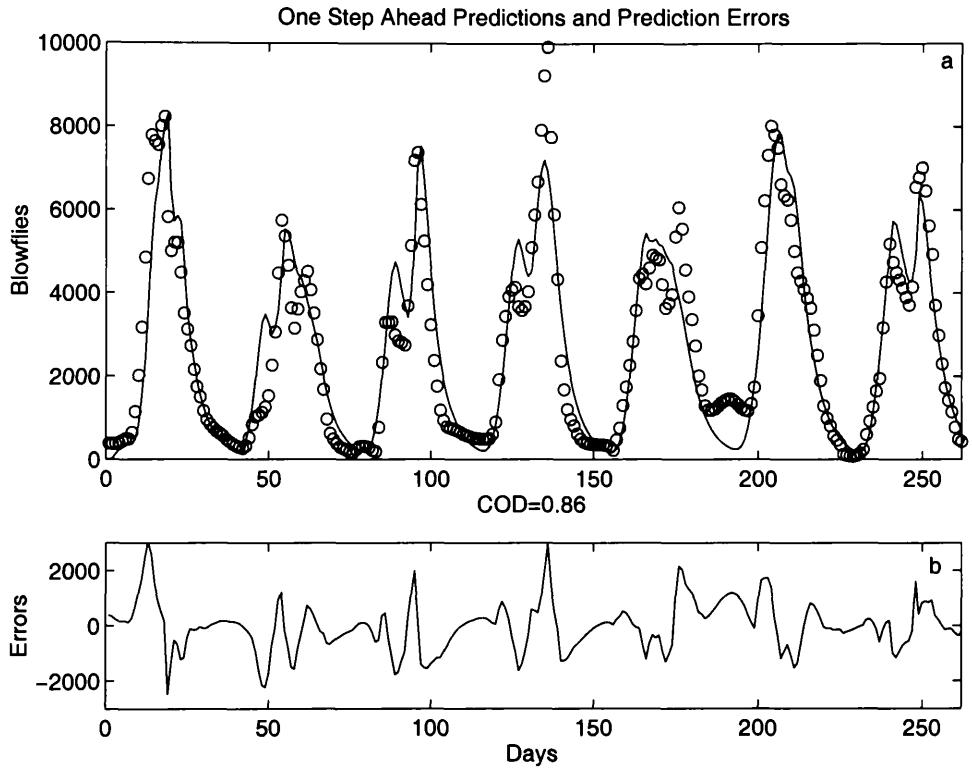

Figure 3.9 (a) ML optimised stochastic one-step ahead predictions (fine line) versus measured blowfly data (circles). (b) Prediction Errors.

As in the deterministic case, the model one-step ahead prediction errors can be analysed using the autocorrelation (AC) and partial autocorrelation (PAC) functions to evaluate whether the series is either a serially uncorrelated sequence of white noise or a sequence of coloured noise with definite structure. Figure 3.10 presents the AC and PAC functions of the model residuals and indicates that the series is, indeed, autocorrelated with a 'coloured' temporal pattern.

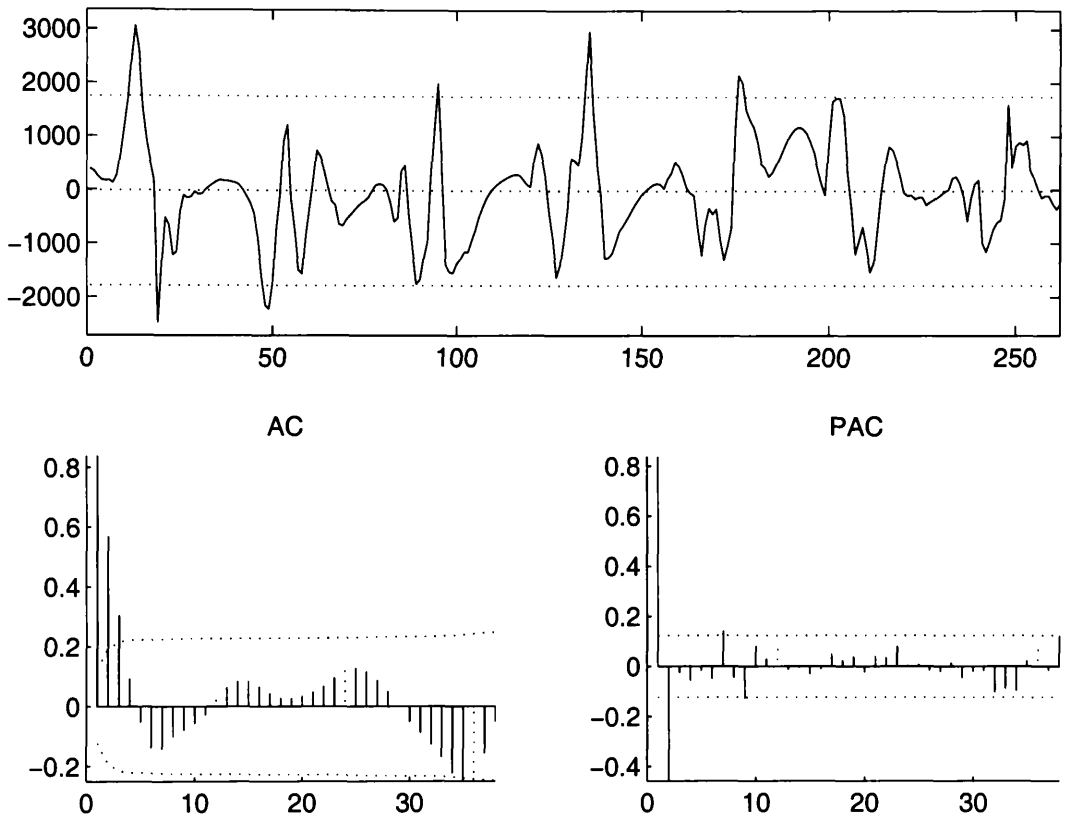

Figure 3.10 One step-ahead prediction errors (top), autocorrelation (bottom left) and partial autocorrelation (bottom right) functions for the stochastic model optimised by ML.

Theoretically, the Kalman Filter can be used for state estimation when the system and measurement noise of the stochastic state space model are independent, zero mean, Gaussian white noise processes. As the residuals of the blowfly state space model in its present form are 'coloured', the state estimates from the Kalman Filter may not be truly optimal and are probably biased from these optimal values.

Moreover, when the optimisation routine is initiated with a high *NVR* value, convergence occurs at a different location in the parameter hyper-space. In this situation, although the model produces much smaller one-step ahead prediction errors, the limit cycle is greatly inferior and, therefore, has both a very poor long-term predictive capability and unacceptable physical meaning. This suggests that a local minimum was initially located to give the parameter values shown in equation (3.20-3.22) and the global minimum is found when the optimisation is started with this higher *NVR* value. However, this global minimum is clearly unsatisfactory and it suggests that the model (3.17-3.19) needs to be modified to account for the stochasticity in the data in order to accommodate smaller one-step ahead prediction errors *and*, at the same time, an acceptable limit cycle that will produce good long term predictions.

3.6 EXTENSION OF THE STOCHASTIC BLOWFLY MODEL

Model (3.17-3.19) can be extended by introducing additional stochastic state variables to characterise the 'coloured' one-step ahead prediction errors observed to be present in Figure 3.10, which suggest a second order stochastic process. By modelling the 'coloured' noise as a second order auto-regressive (AR) process, the model (3.17-3.19) can be extended to form a 17th order stochastic state space model. With this modification, the state, measurement and time varying parameter equations remain unchanged, (equations 3.17 and 3.19) but the state vector $\boldsymbol{x}(k)$, the 17×17 state transition matrix $\boldsymbol{F}$, the observation matrix $\boldsymbol{H}$ and the noise term $\xi(k)$ need to be redefined as,

$$\boldsymbol{x}(k) = [y(k)\ y(k-1) \ldots y(k-13)\ y(k-14)\ \xi(k)\ \xi(k-1)]^T \tag{3.23}$$

$$\boldsymbol{F} = \begin{bmatrix} -a & 0 & 0 & \cdots & 0 & 0 & j(k) & -c_1 & -c_2 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{H} = [1000000000000000]^T$$

$$\xi(k) = -c_1(k)\xi(k-1) - c_2(k)\xi(k-2) + \varepsilon(k) \tag{3.24}$$

where $\varepsilon(k)$ is a serially uncorrelated sequence of white noise with zero mean. An additional *NVR* term Q_2 is introduced, characterising the variance of the system noise. Now, seven parameters in total require optimisation: a_1, g, M, c_1, c_2, Q_1 and Q_2. The parameter estimates are again obtained by minimising the likelihood function, adopting the same method discussed in the previous section. The final optimised model obtained in this manner, takes the following form:

$$u(k) = 4.6383x(k).\exp\left[-\frac{x(k)}{0.5\times 1352.4}\right] \tag{3.25}$$

$$x(k) = 0.8232x(k-1) + u(k-15) + \xi(k) \tag{3.26}$$

$$\xi(k) = 1.1128\xi(k-1) - 0.4589\xi(k-2) + \varepsilon(k) \tag{3.27}$$

$$y(k) = x(k) + e(k) \qquad e(k) = N\{0, \sigma^2\} \tag{3.28}$$

where $\varepsilon(k)$ is $N\{0,45.69\sigma^2\}$.

The one-step ahead predictions from the model (3.25-3.28) fit the data exceptionally well, as shown in Figure 3.11, explaining 97% of the data ($R^2 = 0.969$). As a direct result of the model reformulation, the AC and PAC functions confirm that the one-step ahead prediction errors are serially uncorrelated, white noise. The error series does exhibit heteroscedastic behaviour, which theoretically should be eliminated. However, for present purposes, the model is certainly adequate.

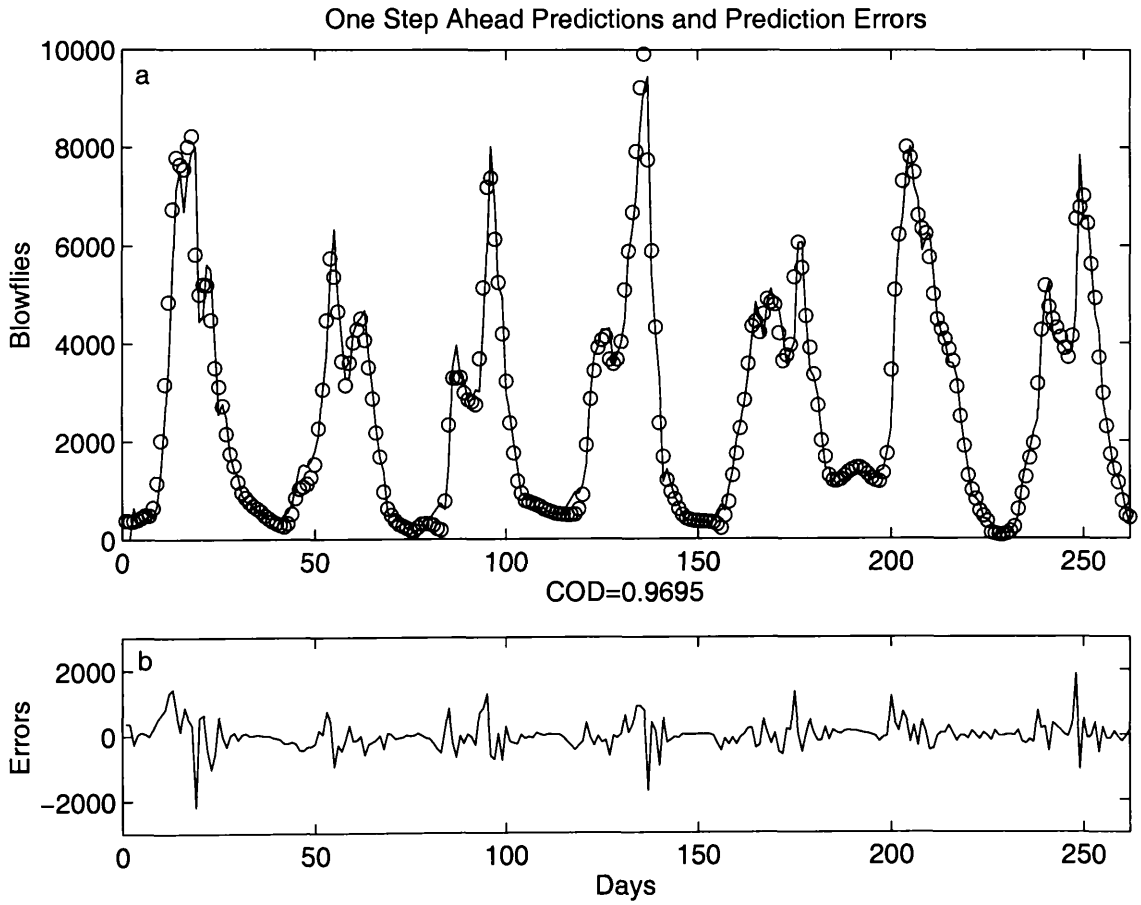

Figure 3.11 (a) One-step ahead predictions (fine line) from the final optimised stochastic model and blowfly time series data (circles) (b) One-step-ahead prediction error series.

As expected, the output from the deterministic component of this model, as shown in Figure 3.12, does not fit the time series data as well as the deterministic model optimised by least squares. There is a simple explanation for this: the stochastic model explains the time series through a combination of the deterministic output and the stochastic output. Therefore, the model parameters are optimised by ML based upon this combined model output. In contrast, by optimising the deterministic model by least squares, the model explains some of the stochastic portion of the time series directly, with the parameters optimised to do this. This is the reason why, the least squares parameter estimates are likely to be biased.

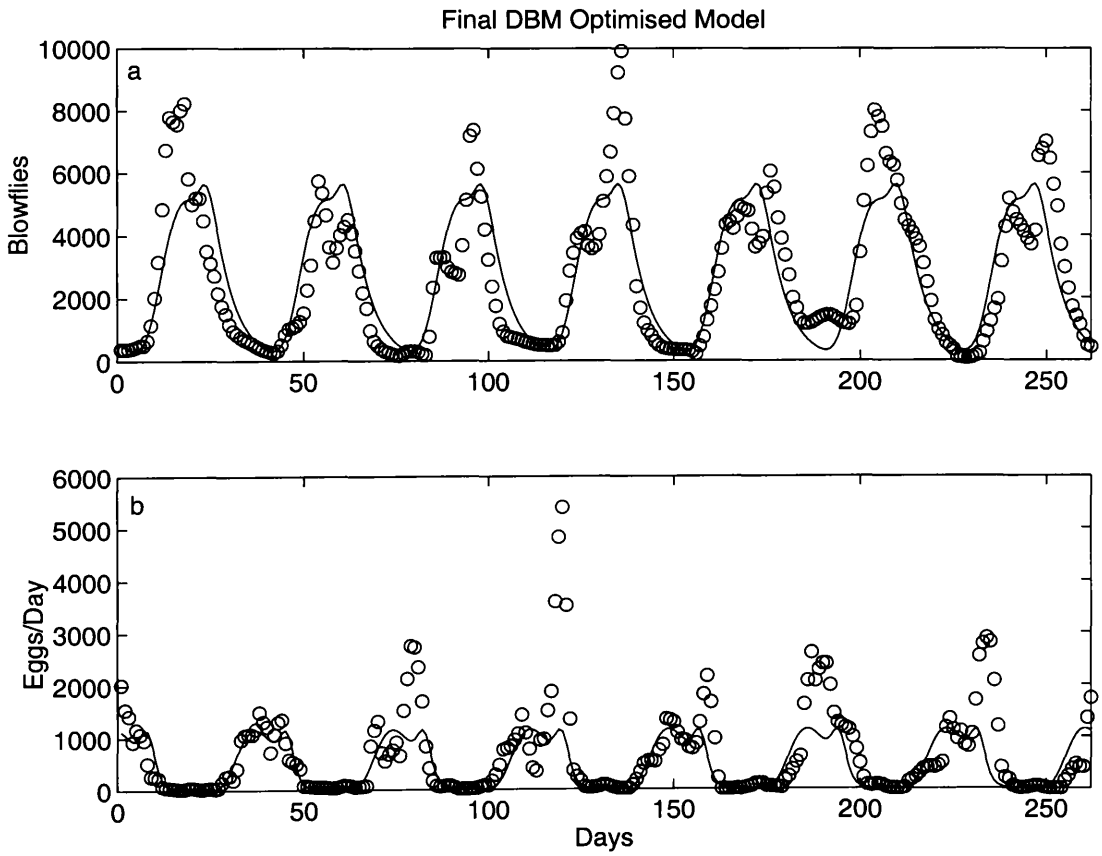

Figure 3.12 The output of the deterministic component of the final stochastic model.

(a) Simulated blowflies (fine line) and data (circles)

(b) Simulated Eggs/Day (fine line) and data (circles).

The stochastic model has two major advantages. Firstly, the predictive ability of the model is far superior. Having formulated the stochastic model in a framework which incorporates the KF, the state and error covariance equations (*prediction equations*) are repeatedly solved to give *a priori* estimates of the blowfly and egg population. Moreover, if additional data should become available, the *correction equations* of the KF can be utilised to correct the predictions and produce *a posterior* estimate of the blowfly and egg population. Likewise, the *a priori* estimate of the error covariance matrix is updated and subsequently the standard error bounds of the predicted blowfly population decrease in width. Figure 3.13 presents the multi-step-ahead predictions initiated at the $151th$ time step. The phase and period of the predicted population limit cycle closely matches the time series. More importantly, the data are mostly encompassed by the standard errors of the stochastic predictions, in contrast to the results outlined in Sections 3.4 and 3.5.

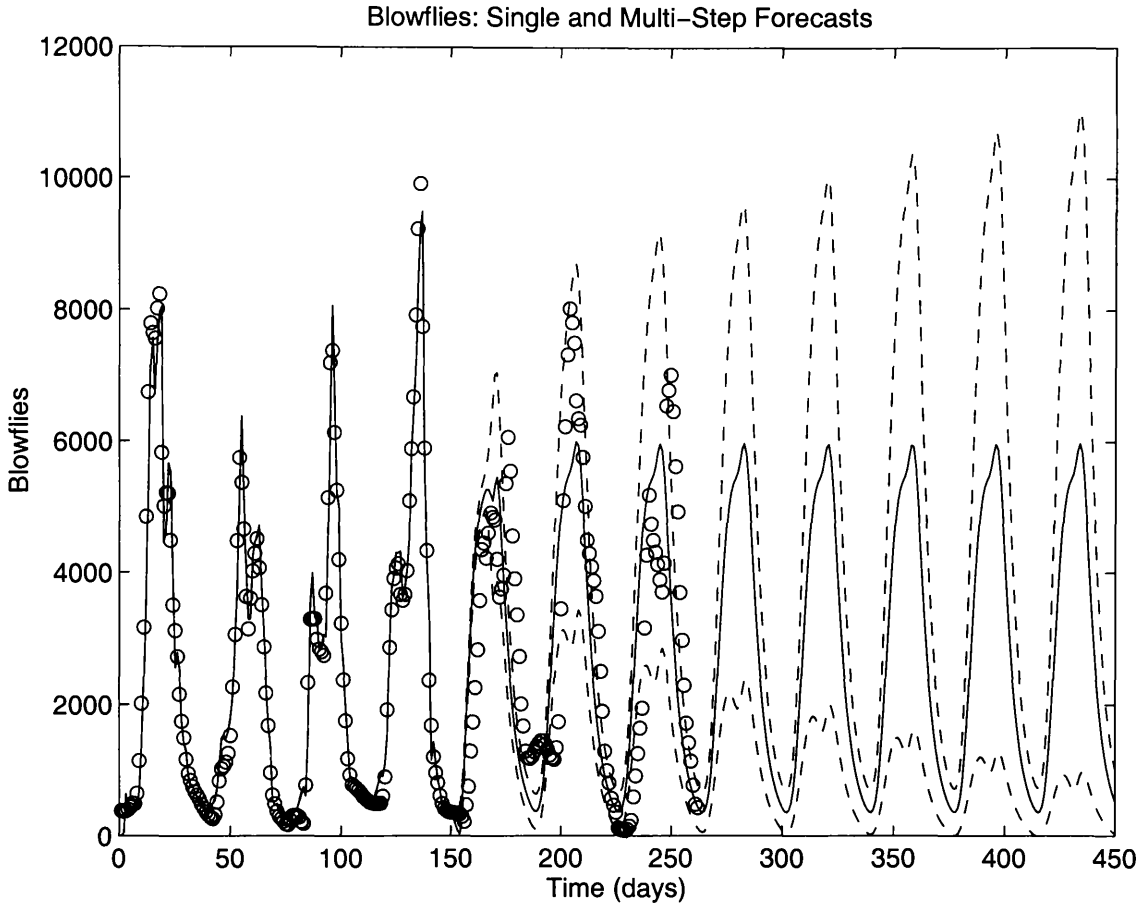

Figure 3.13 Multi-step ahead predictions (fine line), standard errors (dashed line) and blowfly population data (circles).

Secondly, the estimated feedback nonlinearity has an improved, closer fit to the FIS non-parametric estimate, since it now resides well within the standard error bounds, as illustrated in Figure 3.14. In all ways, therefore, the optimised stochastic model (3.25-3.28) satisfies the DBM and statistical requirements of our modelling exercise.

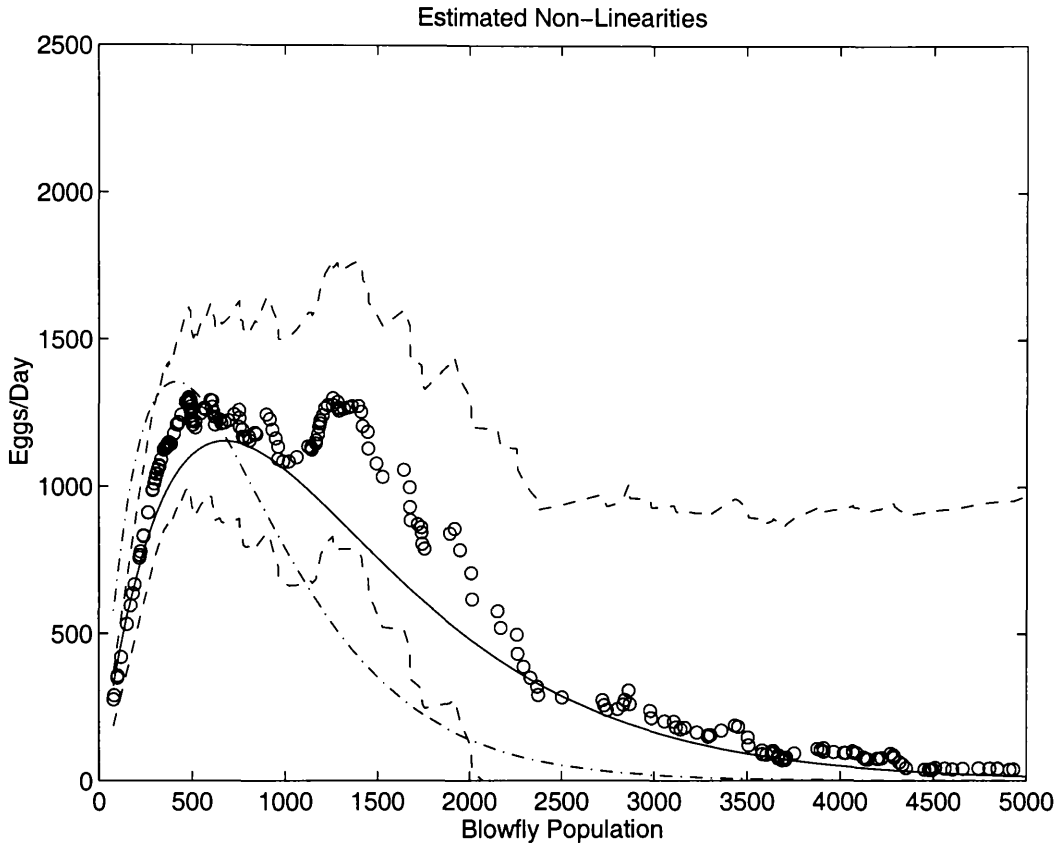

Figure 3.14. Estimated feedback nonlinearities: Non-parametric FIS estimate (circles)

with standard error bands (dashed line), the final stochastically optimised model output (full line) and the deterministic optimised model output (dot dashed line).

Finally, we need to consider model validation. Nominally, this should involve a separate set of experimental data obtained under identical experimental conditions, but this was not possible in this case. Fortunately, Nicholson himself provides one validation possibility. Discussing his experimental results, Nicholson stresses (p 22, 1954) that the blowfly

...'culture was supplied with 0.5 g of ground liver per day and the average density of adults was found to be 2520. In another culture in which all conditions were precisely the same, except only 0.1g of liver was provided per day for adults, the average density of adults was 527.'

The final estimated stochastic nonlinear model can be utilised to make a direct comparison to the average blowfly population observed by Nicholson under the two different environmental conditions. Table 3.2 summaries the average blowfly population predicted from the model (3.25-3.28) and that observed by Nicholson. It is clear that the model comes very close to predicting the observed average blowfly population under both environmental conditions, especially when considering the presence of noise on the measured data. In view of the similarity of these population sizes, the above statement by Nicholson can be considered as an additional, independent, validation of the stochastic model.

Table 3.2 Average blowfly population density deduced by Nicholson (1954) and obtained from the stochastic DBM model

	0.1 g Liver	**0.5 g Liver**
Nicholson	527	2520
Stochastic DBM Model	594	2353

3.7 CONCLUSION

This chapter has outlined a data based mechanistic (DBM) approach to modelling Nicholson's blowfly data. Objective, statistically rigorous procedures have been used to identify and estimate the model structure and parameters directly from the time series data. The forward path of the model (eggs-flies) is characterised by a first order TF but a linear TF fails to describe the relationship between flies and eggs. By adopting a *state dependent parameter* approach, a non-parametric estimate of the feedback relationship has been identified using the *Fixed Interval Smoothing* algorithm. If the DBM model obtained in this manner is formulated in continuous-time, its structure resembles the blowfly model *heuristically* derived by Gurney *et al.*, (1980). Gurney correctly deduced the form of the nonlinear function by visual examination of the data. However, it is very unlikely that modelling, following this 'hypothetico-deductive' approach would necessarily yield a 'correct' model structure in other instances since it depends so much upon the insight of the modeller, which may not always be a good as that of Gurney. In more general terms, the advantages of following a DBM approach to modelling are clearly apparent.

We have found that optimising the model using least squares criteria in a conventional manner fails to produce a satisfactory result: the estimated model is unable to provide acceptable forecasts, the model residuals are 'coloured' and the parameter estimates are likely to exhibit bias. Reformulating the model into a stochastic state space form, allows the model to be optimised in a stochastic manner utilising *Maximum Likelihood.* The resulting one-step-ahead predictions of the optimal model explain the data extremely well and the coloured system noise is successfully modelled by an

additional *auto-regressive* component. Moreover, the predictive ability of the model is superior and the comparison of the ML estimates of the nonlinearity with the non-parametric estimate is much improved.

One clear advantage of the DBM model obtained in this way is that the parameters in the model (3.25-3.28) have a clear ecological interpretation. The parameter a characterises the average daily survival rate of the blowflies; g defines the optimal maximum per capita egg generation rate and N is the population size that produces the maximum reproduction rate. The final model is successfully validated by observations made by Nicholson (1954).

This chapter has also introduced a novel procedure, based on ML, for estimating states in nonlinear models using the Kalman Filter. The nonlinear element of the model is solved and replaced by a time variable (state dependent) parameter at each recursive step before state prediction and correction is undertaken and this allows the Kalman filter to be utilised directly in a 'linear' form. It would appear that this methodology can be applied widely for estimating nonlinear models and this is being investigated in subsequent research studies at Lancaster.

CHAPTER 4

VALIDATION OF A DATA BASED MECHANISTIC NONLINEAR RAINFALL FLOW MODEL

The previous chapter demonstrated how effectively the data-based mechanistic (DBM) modelling approach can be utilised to identify nonlinearities in ecological systems. In this chapter, the DBM methodology is applied to modelling the nonlinear rainfall-flow process (see e.g. Young, 1993). In the absence of extensive field data, rainfall-flow models utilise a surrogate soil moisture series to describe catchment antecedent conditions, the key nonlinearity in the hydrological system. A rainfall-flow model developed at Lancaster (Young, 1993; Young and Beven, 1994) identifies a surrogate soil moisture series directly from the data following the DBM procedures. In contrast, the hybrid data-based conceptual IHACRES rainfall-flow model, applies a conceptual approach to generate the soil moisture surrogate, which is less objective than the DBM methodology (Whitehead and Young, 1975; Jakeman *et al.*, 1990a; Jakeman and Hornberger, 1993). To date, the literature presenting the Lancaster DBM and IHACRES rainfall-flow models has not examined how efficiently the respective surrogate soil moisture component captures the dynamics of the *actual* antecedent

conditions of the catchment. For this reason, this chapter evaluates both models in this regard for the first time, utilising data collected from a Swiss catchment by the Soil Physics Group, ETH, Zurich.

At the catchment scale, hydrological systems are inherently nonlinear, in the sense that the antecedent moisture conditions in the catchment prior to a rainfall event fundamentally affect the resultant stream response. Due to economic and time constraints associated with field data collection, measured hydrological time series are often limited to stream discharge and rainfall; any additional information providing direct measures of soil moisture are the exception rather than the rule. However, for a rainfall-flow model to characterise the nonlinear catchment dynamics well, it is essential that some measure of the antecedent soil moisture condition is incorporated into the model. Typically, therefore, such models account for the antecedent conditions by determining the *effective rainfall* or *rainfall excess* (the total rainfall that directly contributes to the storm runoff) based on a suitable surrogate for soil moisture content. If this approach is successful, then the resulting relationship between effective rainfall and storm runoff is approximately linear and can be modelled in linear transfer function terms.

There are four key modelling objectives in this chapter. Firstly, to introduce the Lancaster DBM and IHACRES rainfall-flow models and to compare their relative performance using two data series obtained from the Swiss experimental catchment. Secondly, to evaluate the effectiveness of the soil moisture surrogates used in the Lancaster DBM and IHACRES models by comparing them directly to field measurements of the *actual* antecedent soil conditions obtained from the Swiss

catchment. Thirdly, to present an extra modification to the time varying parameter (TVP) estimation procedure used by the Lancaster DBM model, which enhances the description of the soil moisture nonlinearity identified from the rainfall-flow time series. And finally, to modify the existing Lancaster DBM model by replacing the soil moisture surrogate with variables representing the *actual* catchment antecedent soil moisture conditions and comparing the new model performance with that of the standard model (c.f. Young, 1993 and Young and Beven, 1994).

4.1 RAINFALL-FLOW MODELLING

Over past decades, key hydrological research has focused on investigating the fundamental relationship between rainfall and river flow and has in turn produced a wide variety of catchment rainfall-flow models. Continual improvements to these models by hydrologists are motivated by the need for increasingly accurate estimates of river flows required for many water resources applications e.g. real time flood forecasting (Lees *et al.*, 1994), water quality analysis (Whitehead and Young, 1975), hydrological impact assessments, and climate change studies. Rainfall-flow modelling techniques can be broadly classified under four major headings: *unit hydrograph*; *conceptual*; *physically based* and *transfer function models*. Each technique is described briefly in the following sections.

4.1.1 Unit hydrograph theory

The concept of unit hydrograph theory (Sherman, 1932) is well known and has been described extensively in the literature (e.g. Shaw, 1984; Wheater *et al.*, 1993). The

unit hydrograph is defined as the stream flow response resulting from a unit input of effective rainfall falling over a catchment in a specified unit of time (the 'unit impulse response' in systems terms). Traditional unit hydrograph estimation involves the subtraction of baseflow (or slowflow) from the stream flow and calculation of effective rainfall, which is then assumed to be linearly related to residual 'quickflow'. In its traditional derivation, the unit hydrograph method has many limitations, such as the arbitrary methods used for base flow separation (see for example Reed, 1984; Shaw, 1984 and Littlewood and Jakeman, 1994), which reduce its practical utility as a rainfall-flow modelling technique.

4.1.2 Conceptual models

The development of powerful computers has enabled hydrologists to design complex rainfall-flow models based on a conceptual representation of the hydrological system. The structure of conceptual models are formulated from simplified representations of the stores and processes that are perceived as important components of the catchment system e.g. evapotranspiration, snow melt, interception storage and groundwater storage. Each component is described by a linear or nonlinear empirical equation and the model is calibrated by optimising the model parameters so that the model output fits the observed data in some optimal (e.g. least squares) sense. Many different conceptual rainfall-flow models have been developed over the last three decades including the well known Stanford Watershed Model (see e.g. Fleming, 1975) and DISPRIN (see e.g. Jamieson and Wilkinson, 1972).

4.1.3 Physically based models

Physically based models are spatially distributed, with the catchment divided into a network of sub-areas or grids. Classical mathematical-physics equations, based on continuum mechanics, are utilised to represent the component processes within the hydrological system for each grid element and are solved through numerical solution using finite element or finite difference methods. Recent examples of physically based models include the Système Hydrologique Européen (SHE) model (Jonach-Clausen, 1979) and the Institute of Hydrology Distributed model (IHDM) (Beven, 1987).

One major disadvantage of conceptual and physically based models is that they often become highly parameterised as a result of representing the physical complexities of the hydrological system. Furthermore, the statistical problems associated with over parameterisation can seriously prejudice the useful application of these models. Recently, researchers (see e.g. Loague and Freeze, 1985; Hornberger *et al.*, 1985; van Gunuchten, 1991; Jakeman and Hornberger, 1993) have recognised that the information present within stream flow and precipitation records does not necessarily warrant the complex parameterisation of these models. Beven (1989, p. 159) reports that '*It appears that 3 to 5 parameters should be sufficient to reproduce most of the information in a hydrological record*'. This compares to the 21 parameters required for the DISPRIN model. For these reasons, some researchers are becoming increasingly interested in developing and applying more parametrically efficient models.

4.1.4 Transfer Function (TF) models

A simpler and more objective approach to rainfall-flow modelling exploits the systems identification and estimation procedures outlined in Chapter 2, where the model structure, including the nature of the soil moisture nonlinearity, is directly inferred from the time series data.

The statistical data-based approach first presented by Box and Jenkins (1970) has been criticised because the subsequent model structure is perceived as having no obvious physical relationship to the system: i.e. it is considered to be a 'black box' model. However, the DBM modelling approach builds on statistical data-based modelling and extends this intrinsically parsimonious approach so that the model parameters and structure are also interpreted in physically meaningful or 'mechanistic' terms. In this regard, it is interesting to note that a rainfall-flow model can be directly related to unit hydrograph (UH) theory. The TF model response to an impulse of effective rainfall can be considered as the underlying UH for the *total* stream flow. The derivation of a UH for *total* stream flow from the TF model has advantages over traditional procedures (Section 4.1.1): the UH is identified directly from the data and is parameterised in an efficient manner. In addition, as baseflow separation is avoided the integrity of the UH for *total* stream flow is only compromised by errors introduced in making assumptions regarding effective rainfall estimation rather than the combination of these errors and those introduced from baseflow separation.

Furthermore, depending on the order of the TF numerator or denominator polynomials, the TF can be decomposed unambiguously into a series and/or parallel connection of subsystems (Young, 1992; Littlewood and Jakeman, 1994). In a

hydrological context these subsystems can be regarded as linear stores or single unit hydrographs, controlling the dynamic partitioning of effective rainfall through to stream flow. The parameters of these subsystems can also be interpreted in physical terms and will be discussed in more detail in Sections 4.3.3-4.3.4.

4.2 TF RAINFALL-FLOW MODELS

The DBM modelling approach was used to develop the Lancaster rainfall-flow model (see e.g. Young, 1993; Young and Beven, 1994), whereby the information within the rainfall and flow data, is used to determine the model structure. The model utilises a constant parameter, Single-Input-Single-Output (SISO), linear discrete-time TF model (although the model could be alternatively derived in either continuous time or by using the delta operator) to describe the transformation of effective rainfall $u_e(k)$ into an estimate $\hat{x}(k)$ of the flow $y(k)$. The general rainfall-flow model comprises two components: the nonlinear effective rainfall term (equation 4.1) and a second linear component relating effective rainfall to flow (equation 4.2), which are defined as follows,

$$u_e(k) = r(k)s(k) \tag{4.1}$$

$$\hat{x}(k) = \frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})} u_e(k-\delta) \tag{4.2}$$

$$y(k) = \hat{x}(k) + \zeta(k) \tag{4.3}$$

where $r(k)$ is measured rainfall; $s(k)$ is a suitably defined soil moisture surrogate; δ is the pure time delay, $\zeta(k)$ is a general noise term included to account for stochastic

disturbances and unmeasured inputs to the system. The polynomials $\hat{A}(z^{-1})$ and $\hat{B}(z^{-1})$ are defined as,

$$\begin{aligned}\hat{A}(z^{-1}) &= 1 + a_1 z^{-1} + \ldots + a_n z^{-n}; \\ \hat{B}(z^{-1}) &= b_0 + b_1 z^{-1} + \ldots + b_m z^{-m}.\end{aligned} \tag{4.4}$$

Here, z^{-1} is the backward shift operator where $y(k)z^{-i} = y(k-i)$; while the integers n and m are the number of parameters in the model (denoted by the triad $[n,m,\delta]$) which is defined by a method of statistical model order identification.

The hybrid conceptual data-based IHACRES rainfall-flow model also utilises a TF to describe the transformation of effective rainfall to flow and also exploits system identification techniques to determine its structure and parameterisation. Both models have provided very good results in a variety of different conditions, including catchments in the UK, Australia and the USA (Jakeman *et al.* 1993; Young and Beven, 1994; Young *et al.*, 1998). The key difference between the Lancaster DBM and IHACRES rainfall-flow models is in the formulation and estimation of the nonlinear effective rainfall input.

4.2.1 Lancaster DBM Rainfall-Flow Model

In adopting the DBM philosophy, no prior assumptions about the possible nature of the soil moisture nonlinearity are made in the rainfall-flow model of Young (1993). Rather, this nonlinearity is objectively identified from the measured rainfall-flow

series using *time varying parameter* (TVP) estimation based on *Fixed Interval Smoothing* (FIS), discussed in Chapter 2.

As the rainfall-flow system is inherently nonlinear, a linear model will not normally explain the data well, except for isolated storms and during extremely wet conditions. In more general situations, statistical tests on the residuals of a linear model indicate the presence of a nonlinearity due to the soil moisture effects. The only way in which the model can adequately explain the data, is either to explicitly include nonlinear terms (e.g. IHACRES conceptual nonlinear loss module discussed in Section 4.2.2) or to allow the model parameters to vary over time, so reflecting the changing dynamics induced by the nonlinearity.

A TVP version of the linear TF model (equation. 4.1-4.3) can be written in the following vector form:

$$y(k) = \boldsymbol{z}(k)^T \boldsymbol{a}(k) + \eta(k); \tag{4.5}$$

where,

$$\begin{aligned} \boldsymbol{z}(k)^T &= [-y(k-1) - \ldots - y(k-n)\, r(k) \ldots r(k-m)], \\ \boldsymbol{a}(k) &= [a_1(k)\, a_2(k) \ldots a_n(k) b_0(k) \ldots b_m(k)]^T, \end{aligned} \tag{4.6}$$

and $\eta(k)$ is a stochastic white noise term. The FIS method of recursive estimation described in Chapter 2 can be used to obtain an off-line, non-parametric estimate $\hat{\boldsymbol{a}}(k|N)$ of the parameter vector $\boldsymbol{a}(k)$ in this model, where N is the total sample size. Any parameter that is found to be significantly time variant over the interval N can be

investigated more thoroughly by holding all other parameters constant during FIS estimation.

In rainfall-flow systems, the cause of the time variability is likely to be due to either of the following reasons: slow variations which reflect the changing nature of the catchment due to natural or anthropogenic changes e.g. natural or enforced changes in land use; or more rapid variations caused by nonlinear behaviour. In this latter case, it is likely that the temporal changes are actually caused by some underlying state dependency e.g. due to evapotranspiration and antecedent soil moisture effects.

If any significant parameter variations exhibit state dependence, then it is often possible to establish a relationship between the TVP vector $\hat{\boldsymbol{a}}(k|N)$ and one or more state variables of the system. In the case of rainfall-flow systems, the TVP estimate $\hat{\boldsymbol{a}}(k|N)$ might logically be expected to be a function of variables such as soil moisture and evapotranspiration, i.e.,

$$\hat{\boldsymbol{a}}(k|N) = \alpha \boldsymbol{M}(k) + \varepsilon(k) \tag{4.7}$$

where,

$$\boldsymbol{M}(k) = f(\text{evaporation, soil moisture content etc}) \tag{4.8}$$

By exploiting the covariance matrix $\boldsymbol{P}^*(k|N)$, derived from the FIS algorithm, α can be estimated by minimising the *weighted least squares* (WLS) cost function as defined previously (equation. 2.38).

Experience with rainfall-flow data has shown that the parameters of the $\hat{A}(z^{-1})$ denominator polynomial are normally time invariant, while those of the $\hat{B}(z^{-1})$ numerator polynomial tend to show significant variation over the sampling period, suggesting the presence of an input nonlinearity. Thus, in the simplest of cases where n and m in equation (4.5) are both unity, only one single TVP, $\hat{b}_0(k|N)$ needs to be estimated. Furthermore, it has been established that the TVP $\hat{b}_0(k|N)$, relates well to a function of delayed flow $y(k-\delta)$, (Young, 1993). Both, two-stage-linear and power law relationships between $\hat{b}_0(k|N)$ and $y(k-\delta)$ have been reported (Young, 1993).

A power law relationship describing catchment antecedent conditions is in common with other research of a more physical nature, (Lynch *et al.,* 1979; Myrabo, 1986). Moreover, since a catchment can be seen as a low pass filter, continually smoothing the intense fluctuating rainfall to generate stream flow, it also seems reasonable that stream flow in turn, will reflect the antecedent rainfall and hence soil moisture conditions of the catchment. For example, when the flow is high it suggests the antecedent conditions were wet and the soil water deficit is low. Subsequent effects of rainfall on stream flow will, therefore, be large as more runoff is generated. An effective rainfall series $u_e(k)$ can be generated from the product of the state dependent relationship and the measured rainfall (equations 4.9-4.10), where the scaling parameter vector α has been eliminated since it can be absorbed into the TF model.

$$u_e(k) = r(k)s(k) \tag{4.9}$$

$$s(k) = y(k)^{\beta} \tag{4.10}$$

4.2.2 IHACRES rainfall flow model

In contrast to the Lancaster DBM model, the IHACRES model (see e.g. Jakeman and Hornberger 1993) utilises a conceptually formulated nonlinear module, which exploits a catchment wetness index derived from past rainfall and where appropriate, temperature data, to obtain effective rainfall. The IHACRES model is a slightly modified version of the rainfall-flow model first developed by Young and Whitehead from studies modelling the Bedford-Ouse River (Young, 1974; Whitehead and Young, 1975; Whitehead *et al.*, 1979; and Young 1984). Young and Whitehead introduced a conceptual 'rain filter' to derive effective rainfall which incorporates the effects of soil moisture and temperature dependent evapotranspiration, encapsulating the theory that rainfall, falling on a wet catchment, will generate a larger flow than if the catchment is dry. The 'rain filter' first adjusts rainfall $r(k)$ to account for evapotranspiration losses using a temperature dependent factor. The adjusted rainfall series $r^*(k)$ is then multiplied by a running catchment wetness index $s(k)$, which itself, is derived from past rainfall (and temperature) data. Although data from a number of catchments have been successfully modelled utilising this 'rain filter' (see e.g. Jakeman *et al.*, 1990a; 1993), it does, however, have a conceptual weakness: in the absence of rainfall, the 'rain-filter' does not directly allow for any evapotranspiration losses from the catchment. This shortcoming has led to the derivation of a number of more conceptually acceptable rain-filters. The basic IHACRES nonlinear loss module calculates effective rainfall $u_e(k)$ by multiplying the catchment wetness index $s(k)$ by the measured rainfall $r(k)$ as defined by equations (4.11-4.14),

$$u_e(k) = s(k)r(k) \tag{4.11}$$

$$s(k) = cr(k) + (1 - 1/T_w)\, s(k-1) \tag{4.12}$$

The catchment wetness index (i.e. soil moisture content) $s(k)$ is obtained at each time step k $(k = 1,2,\ldots,N)$ from an exponentially decaying weight of rainfall $r(k)$ at previous time instants. The decreasing influence of past rainfall episodes on $s(k)$ can be clearly shown in the full expansion of equation (4.12), where the term $(1-T_w^{-1})^N$ gets smaller with time,

$$s(k) = c\left[r(k) + (1-T_w^{-1})r(k-1) + (1-T_w^{-1})^2 r(k-2) + \cdots + (1-T_w^{-1})^N r(k-N)\right] \tag{4.13}$$

Parameter T_w is a time constant representing the decay rate of the catchment wetness (or soil moisture) in the absence of rainfall. The soil properties of the catchment are controlled by this parameter; the lower the value of T_w the faster the catchment responds to the processes of wetting and drying and vice versa. Furthermore, in climates where evapotranspiration rates significantly influence the catchment wetness dynamics, the coefficient T_w, can be assumed to vary as a function of temperature $t(k)$.

$$T_w\big(t(k)\big) - T_w \exp[20 - t(k)f] \tag{4.14}$$

$T_w(t(k))$ is, therefore, inversely related to the decline rate of the catchment wetness at $20°\mathrm{C}$, modulated by parameter f. The values of parameters T_w and f are obtained through objective optimisation, but the value of parameter c is selected such that the total volume of effective rainfall is equal to the total volume of discharge over the calibration period. For short time series, where changes in temperature are not of

sufficient magnitude to cause any significant evapotranspiration effects, this extra term $T_w(t(k))$ can be held constant.

Additional conceptualised versions of this nonlinear loss module have been designed to include additional complexities, for example, to account for interception of incident precipitation by tree cover (Jakeman *et al.* 1994; Chen *et al.*, 1995). However, the module described by equations (4.11-4.14) is adequate for the purposes of the present study.

4.3 MODEL CALIBRATION AND VALIDATION

In the research to date, the efficiency of the effective rainfall measure $u_e(k)$ in both the Lancaster DBM and the IHACRES models has only been evaluated in relation to the performance of the models as a whole. In this section, this aspect of the model is evaluated more directly by comparing the surrogate soil moisture measures with the actual antecedent dynamics measured at the Swiss catchment.

4.3.1 The Data

The data used in this research has been collected from the narrow Erienbach catchment situated in the Swiss pre-Alps by the Soil Physics Group at ETH, Zurich, as part of the on going NITREX project studying the effects of nitrogen addition to small catchments (Wright *et al.*, 1995). Erienbach is situated at 1200m a.s.l with a total area of 0.7km^2 of which approximately 40% is forested and 60% is wetland and has an average total yearly precipitation of 2300mm.

Two data series (Data Series 1 and 2) have been collected from a muck humus soil plot with an approximate area of $15m^2$, representing one of two soil types which characterise the catchment. The input and state variables measured at the plot scale are, rainfall $r(k)$; flow $y(k)$; percentage soil water content $pwc(k)$; and ground water table depth $gw(k)$. Where $pwc(k)$ and $gw(k)$ were derived from measurements made by sixteen soil moisture probes (TDR) and three piezometers respectively. A photograph of the soil moisture probes *in situ* is shown in Figure 4.1. The sampling interval for each state was 10 minutes although average hourly readings have been used in the analyses reported in this chapter.

Figure 4.1 Soil moisture probes *in situ* at the Erienbach catchment

Figure 4.2 shows the input and state variables from Data Series 1. The immediate response to rainfall inputs can be observed in both soil water measurements $pwc(k)$ and $gw(k)$. Similar responses are observed in Data Series 2 as shown in Figure 4.3.

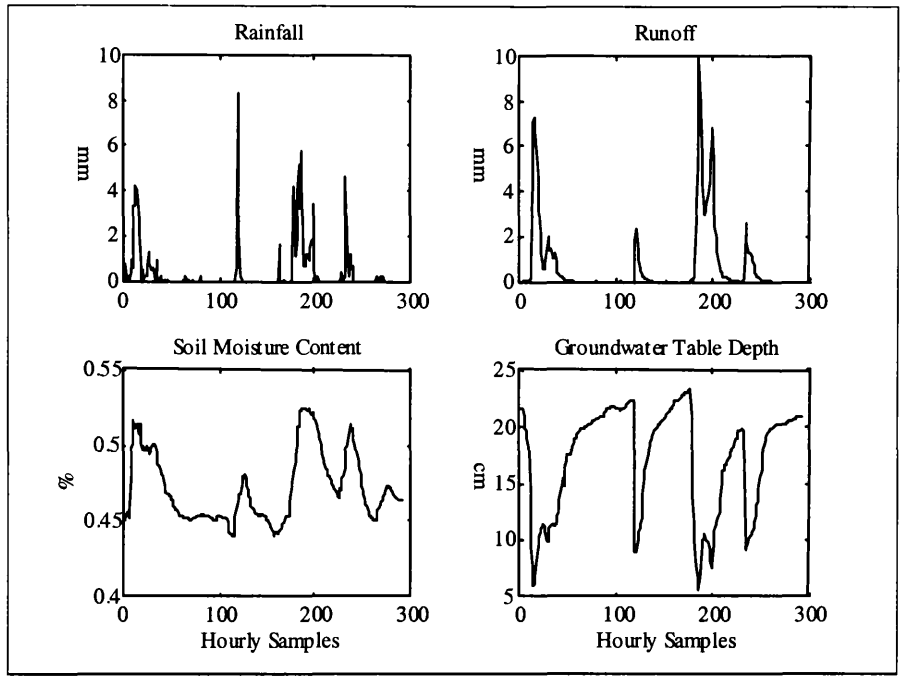

Figure 4.2. Measured Variables, Data Series 1.

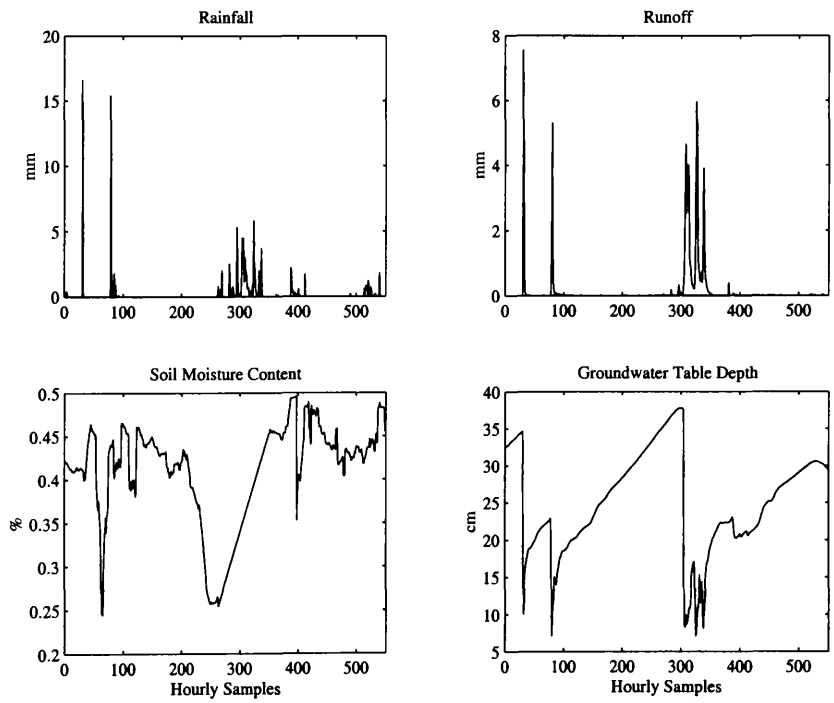

Figure 4.3. Measured Variables, Data Series 2.

There is, however, some doubt over the quality of the soil moisture content data in Data Series 2. The timing and the dynamics of this time series do not appear consistent with rainfall, flow and depth to groundwater table data. Episodes of rainfall result in a decrease in soil moisture content which clearly, makes no physical sense. Consultation with H. Feyen (*pers. com.*) from the Soil Physics Group, ETH, highlighted that the soil moisture content monitoring equipment had malfunctioned during this time period. Consequently, this percentage soil moisture content time series (Data Series 2) was disregarded for use within the modelling study. The next two sections utilise the Erienbach catchment data to identify and estimate both the Lancaster DBM and IHACRES nonlinear rainfall-flow models. The identified surrogate soil moisture series is then compared to the percentage soil moisture content $pwc(k)$ and depth to groundwater table $gw(k)$ variables.

4.3.2 Linear transfer function model

Visual examination of Data Series 1 clearly shows nonlinear behaviour existing between the rainfall and flow: similar sized rainfall events produce different flow effects depending upon the prevailing antecedent conditions. The SRIV algorithm identifies the best linear, *constant parameter* TF model between measured rainfall $r(k)$ and flow $y(k)$ as a first order [1,1,0] structure, with model statistics of $R_T^2 = 0.845$ and $YIC = -8.049$. Figure 4.4 shows how the linear model is unable to capture the dynamics of the flow, particularly during the receding limb of the hydrograph. The following two sections demonstrate how this model can be improved by incorporating a nonlinear effective rainfall term.

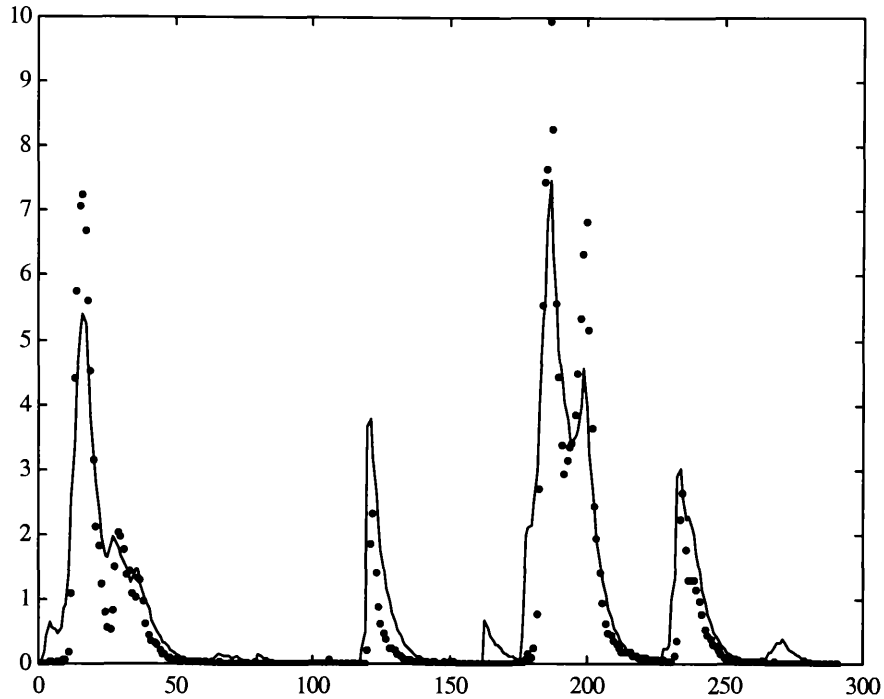

Figure 4.4 Best *constant parameter* linear TF model between rainfall and flow. Model (line) and measured data (dots).

4.3.3 Lancaster DBM model results

On examination, the residuals of the first order linear TF model identified in the previous section are, as expected, heavily structured and confirm that the data are being affected by a nonlinearity. TVP estimates from the FIS algorithm (following NVR optimisation) suggest that only $\hat{b}_0(k|N)$ is varying significantly with substantial temporal variation, as shown in Figure 4.5.

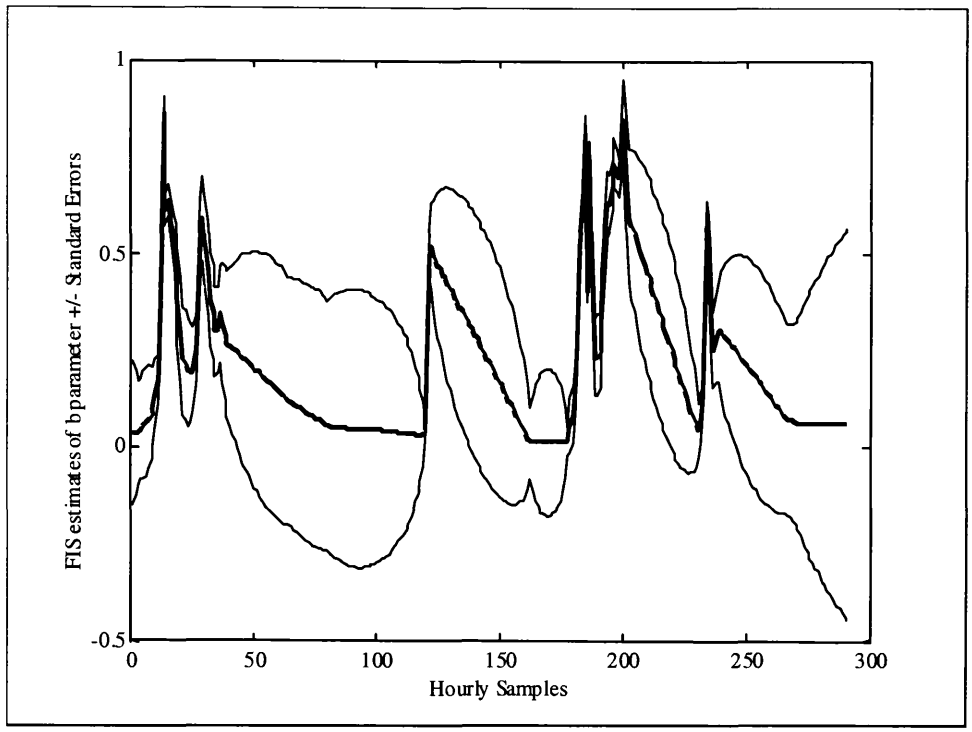

Figure 4.5. FIS estimate of TVP $\hat{b}_0(k|N)$ (bold line) and standard error bands (fine line), Data Series 1.

The similarity between $\hat{b}_0(k|N)$ and the flow $y(k)$ (Figures 4.2 and 4.5) is visually apparent and suggests that the temporal variation in the parameter can be explained by a state dependent relationship. This nonlinear relationship was identified by the procedure outlined in Section 4.2.1 and the following estimate of the power law function is obtained by WLS estimation and shown in Figure 4.6. In the WLS estimation process, those elements of the normalised weighting matrix that have a value lower than 0.2 are set to zero.

$$\hat{b}_0(k|N) = 0.72 y(k)^{\beta}; \quad \beta = 0.43 \tag{4.15}$$

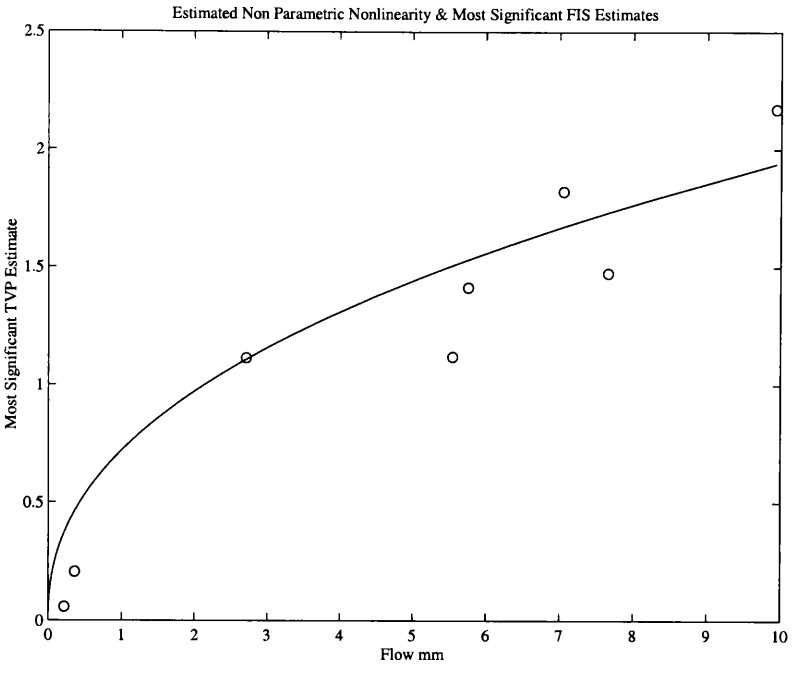

Figure 4.6 The most significant estimates of $\hat{b}_0(k|N)$, shown as circles versus flow and the WLS estimate of the power law relationship.

However, this initial estimate of the power law parameter beta, is obtained from the non-parametric FIS estimate of $\hat{b}_0(k|N)$ and is intended only to *identify* the nature of the nonlinearity. Consistent estimates of the power law relationship are obtained by an iterative optimisation procedure using the SRIV algorithm. The best five nonlinear TF models identified are shown in Table 4.1.

Table 4.1 Best 5 DBM nonlinear TF models ordered in terms of *YIC*

Den	Num	Delay	Beta	YIC	R_T^2	AIC
1	1	0	0.36	-10.585	0.954	-2.068
2	2	0	0.35	-7.479-	0.959	-2.1511
1	3	0	0.36	-6.822	0.966	-2.337
1	2	0	0.32	-6.301	0.957	-2.111
2	3	0	0.35	-6.153	0.972	-2.509

After careful consideration, the results suggest that the nonlinear model with a [1,3,0] structure gives the best compromise between fit to the data and parametric efficiency.

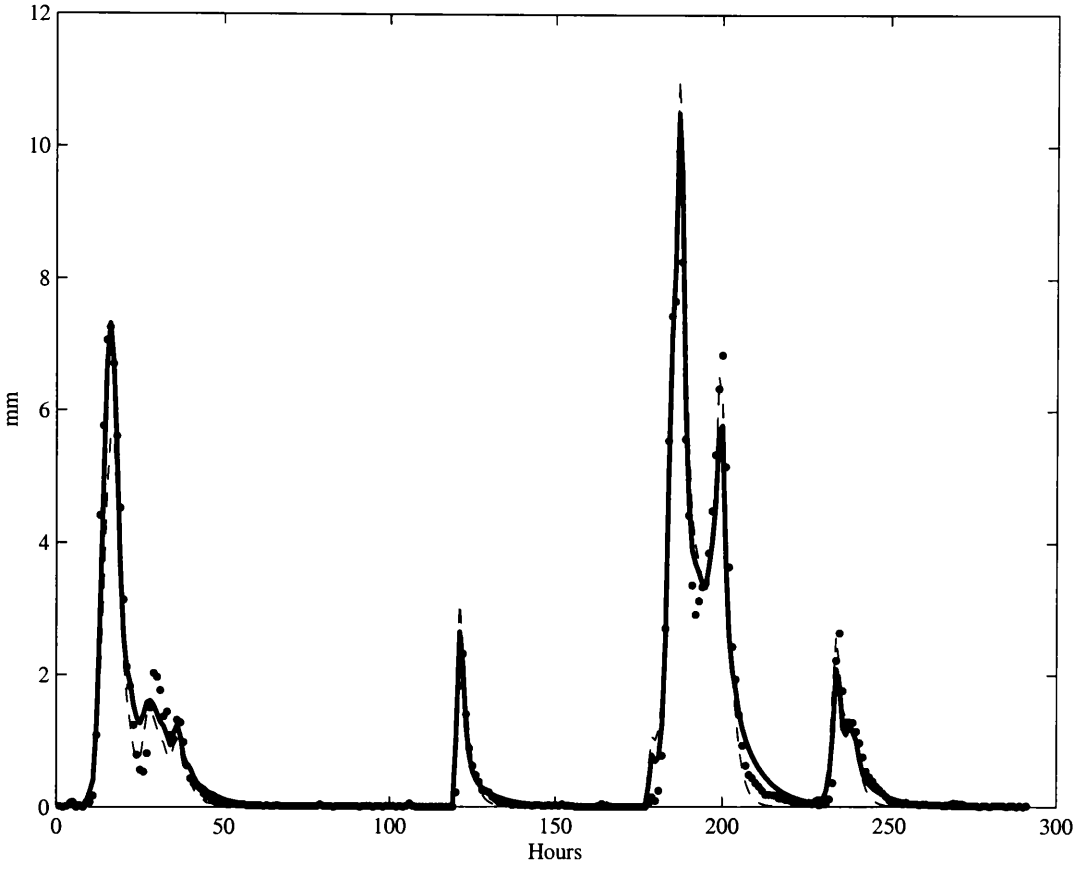

Figure 4.7 The [1,3,0] Lancaster DBM model (full line), [1,2,0] IHACRES model (dashed line) and the observed data (dots).

On the basis of the *YIC* criteria only, the best nonlinear model has a [1,1,0] structure. However, the *YIC* criteria often favours low order models and as the R_T^2 value of the [1,1,0] model is low, it can be rejected since it does not adequately explain the data. The model with the next most negative *YIC* has a [2,2,0] structure. However, the eigenvalues of the denominator polynomial are complex and so, within the context of the DBM modelling philosophy, the model can be rejected as there is no logical physical explanation of this behaviour. The [1,2,0] model can also be disregarded, as

the associated *YIC*, R_T^2 and *AIC* values are inferior to those of the [1,3,0] model. Finally, although the [2,3,0] model has a slightly higher R_T^2 than the [1,3,0] model, from visual inspection of the model output, it is apparent that the [2,3,0] model consistently fails to characterise the recessions of the hydrograph. For this reason and the additionally higher *YIC* value, the [2,3,0] model can also be rejected. The SRIV estimated parameters and associated standard errors of the favoured [1,3,0] model are as follows:

$$a_1 = -0.812(0.010) \qquad b_0 = 0.210(0.019) \qquad b_1 = 0.322(0.036)$$
$$b_2 = -0.256(0.019) \qquad \beta = 0.358(0.007)$$

The excellent [1,3,0] model fit to the data is shown in Figure 4.7. Moreover, the linear TF of the [1,3,0] model can be decomposed into three parallel processes, which have clear physical interpretations. The partitioning of the effective rainfall at the plot scale is shown in Figure 4.8 and takes the following form: an 'instantaneous' component transferring 14.9% of the flow; a 'quick' one hour delayed instantaneous component transferring 34.1% of the flow; and finally, a 'slow' pathway accounting for the remaining 51% of the flow.

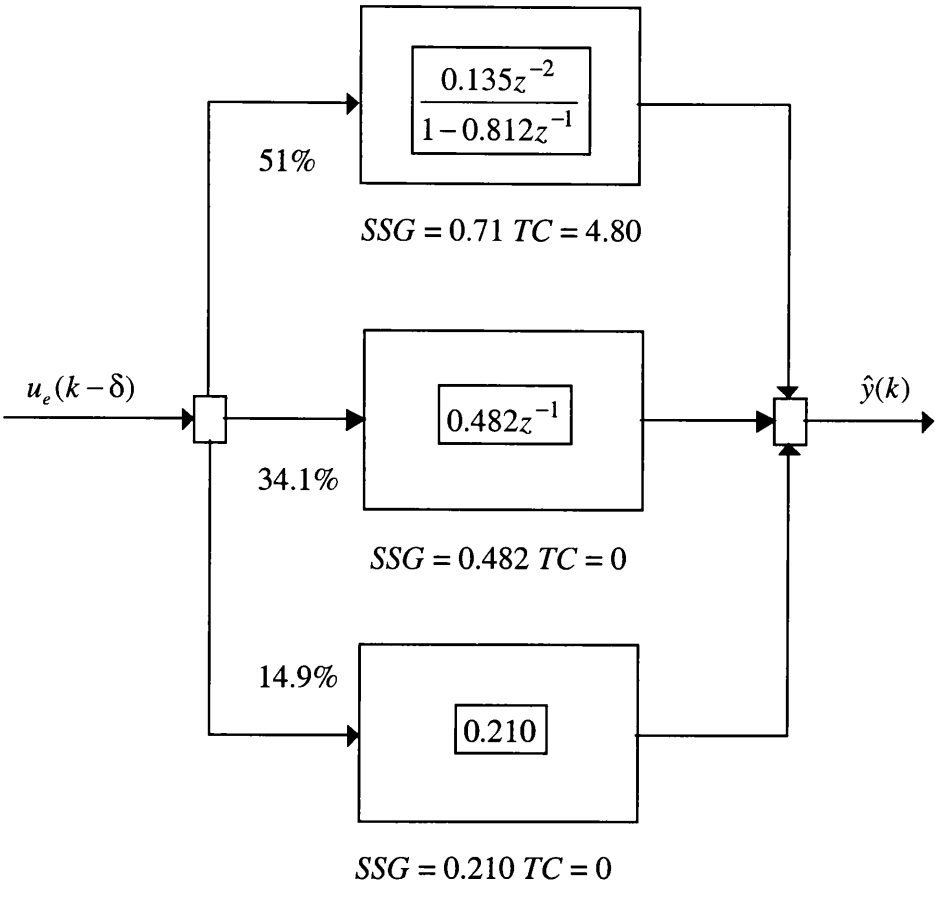

Figure 4.8 Parallel partitioning of flow processes. Where SSG denotes the steady state gain and TC the time constant or residence time.

This physical interpretation is in good agreement with observations and other modelling studies of flow processes at both the soil plot and catchment scale at Erienbach (Feyen *et al.*, 1994). At the plot scale, surface runoff and subsurface flow are measured at three soil depths (5, 30 and 60 cm). The modelling results approximate to the proportion of the total flow and residence times observed at each of the three depths (Feyen, *Pers. Comm.*).

4.3.4 IHACRES model results

To investigate whether a better model fit could be achieved using the nonlinear IHACRES model, an iterative routine, incorporating the SRIV algorithm, was designed to optimise both the parameters in the nonlinear (equations 4.11-4.14) and linear components of the model (equations 4.1-4.4). The parameter $T_w(t(k))$ is held constant as evapotranspiration affects are assumed to be negligible over the short duration of the time series. The best five models and their respective diagnostic fitting criterion are shown in Table 4.2. It is immediately apparent that the IHACRES model gives a reduced model fit when compared with the Lancaster DBM model.

Table 4.2 Best five IHACRES nonlinear TF models ordered in terms of *YIC*

Den	Num	Delay	T_w	c	YIC	R_T^2	AIC
1	1	0	15.061	0.123	-9.189	0.919	-1.480
1	2	0	16.418	0.118	-6.111	0.929	-1.613
2	1	0	16.736	0.117	-5.931	0.926	-1.568
3	1	0	15.607	0.121	-5.557	0.930	-1.619
2	3	0	14.277	0.086	-3.145	0.932	-1.645

Discarding the inferior [1,1,0] model, the fitting criterian show that there is little to distinguish between the remaining models. Although the [3,1,0] and [2,3,0] models do have a slightly better fit (R_T^2) to the data, it is so marginal that, in terms of parametric efficiency, the [1,2,0] model is chosen as the best overall model (see Figure 4.7). The parameters for this model and their standard errors are as follows:

$$b_0 = 0.182(0.283\times10^{-3}) \quad b_1 = 0.157(0.548\times10^{-3}) \quad a_1 = -0.639(0.232\times10^{-3})$$

$$T_w = 16.418(0.377) \quad c = 0.118$$

The linear component of the [1,2,0] model has real eigenvalues and by partial fraction decomposition, can be represented by two flow processes connected in parallel, as shown in Figure 4.9. Of the total effective rainfall, 80.6% is transferred through the plot by means of a 'slow' pathway with a time constant of 2.23 hours with the remaining 19.4% of the effective rainfall travelling through the plot instantaneously.

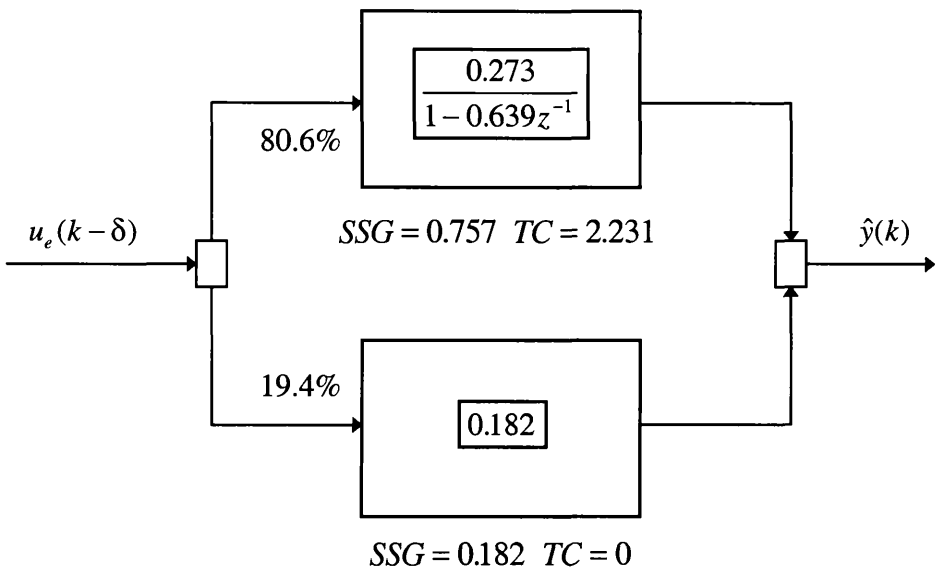

Figure 4.9 Parallel partitioning of the flow process. Where SSG denotes the steady state gain and TC the time constant or residence time.

4.3.5 Monte-Carlo Analysis (MCA)

Whilst the rainfall-flow models identified in Sections 4.3.3-4.3.4 provide good evidence for parallel flow processes, it is important to consider the calculated values of the partitioning coefficients and any associated physical deductions with caution, until some form of model uncertainty analysis is undertaken, such as Monte-Carlo analysis (see e.g. Whitehead and Young, 1979) and further experimentation at the catchment is completed. The estimated parameter variance and covariance associated with the stochastic TF model, can be utilised within the context of Monte-Carlo analysis, to highlight the sensitivity of the parallel partitioning to parametric uncertainty. Monte-Carlo analysis will be discussed in detail in Chapter 6; however, it is sufficient to say that the rainfall-flow model can be simulated repeatedly with parameters taken from probability distribution functions (pdf's) which are defined utilising the parameter covariance matrix obtained from the initial SRIV estimation.

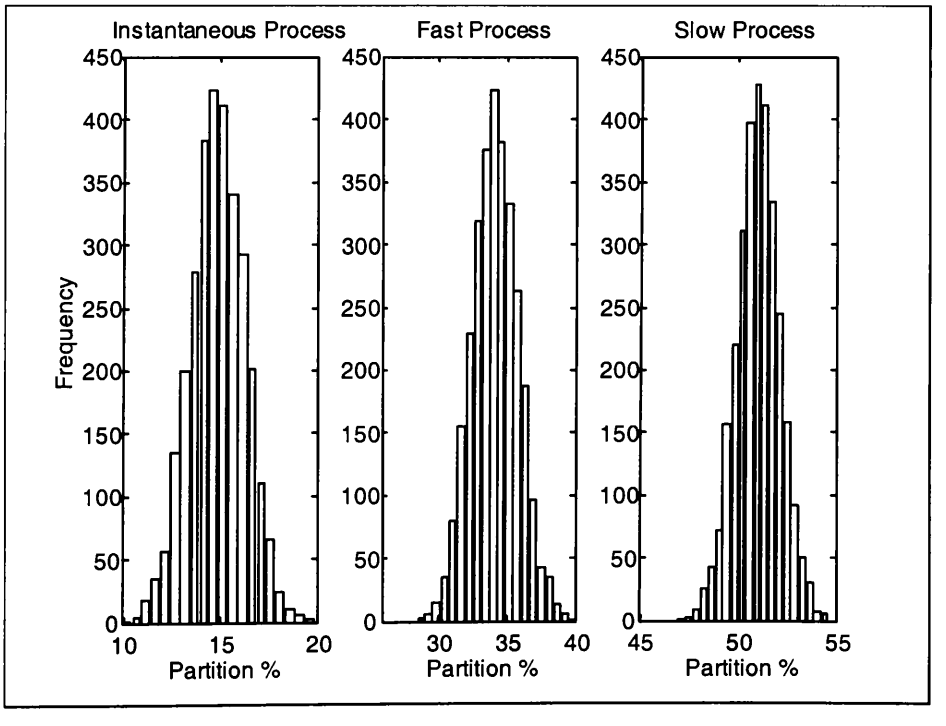

Figure 4.10 Parallel partitioning histogram.

Figure 4.10 presents histograms of the partitioning coefficients, obtained from 3000 simulations of the Lancaster DBM rainfall-flow model identified in Section 4.3.3. The histograms show that each partitioning coefficient is very well defined with mean values for the instantaneous, slow, and fast components calculated as 14.85%(1.40), 34.13%(1.67) and 51.02%(1.10) respectively, with standard errors given in parenthesis. Whilst these results indicate that the partitioning of flow is well defined, in other studies, Monte-Carlo analysis has identified quite the opposite (Young, 1992). In this regard, through the provision of additional information which would otherwise remain unknown, Monte-Carlo analysis should be viewed as a valuable, complimentary tool to modelling and the analysis of environmental systems.

4.4 COMPARISON OF THE MEASURED SOIL MOISTURE VARIABLES WITH GENERATED SURROGATE SOIL MOISTURE SERIES

Following on from the identification and estimation of the IHACRES and Lancaster DBM nonlinear TF models (Sections 4.3.3 and 4.3.4), this section is specifically concerned with the comparison between the estimated soil moisture surrogates, $ihsm(k)$ and $lsm(k)$ generated from the two models respectively, and the measured soil moisture variables $pwc(k)$ and $gw(k)$.

4.4.1 Comparison of the measured soil water variables with the Lancaster DBM soil moisture surrogate

The identified Lancaster DBM model defines a soil moisture surrogate $lsm(k)$ as:

$$\begin{aligned} lsm(k) &= y(k)^{\beta} \\ &= y(k)^{0.354} \end{aligned} \tag{4.16}$$

and from the IHACRES model, the soil moisture surrogate $ihsm(k)$ is defined as

$$\begin{aligned} ihsm(k) &= s(k) \\ &= 0.118r(k) + (1 - 1/16.418)ihsm(k-1) \end{aligned} \tag{4.17}$$

Figure 4.11 presents the soil moisture surrogate $lsm(k)$, percentage soil moisture content $pwc(k)$ and normalised/inversed depth to groundwater table $gw(k)$ for Data Series 1. The close association between $lsm(k)$ and both measured soil water variables, $pwc(k)$ and $gw(k)$, is clear and highlights how effective the TVP and power law technique is at identifying and estimating the soil moisture non-linearity. It will be noted that the greatest correlation exists between $gw(k)$ and $lsm(k)$, in particular, the characteristic double-peaked storm responses seen in $gw(k)$ are clearly represented in $lsm(k)$. In general, the peak responses of $gw(k)$ and $lsm(k)$ are proportional and the recessions also have similar shapes. Although $lsm(k)$ does not show as high a correlation to the percentage soil moisture content time series $pwc(k)$, the overall dynamics are similar.

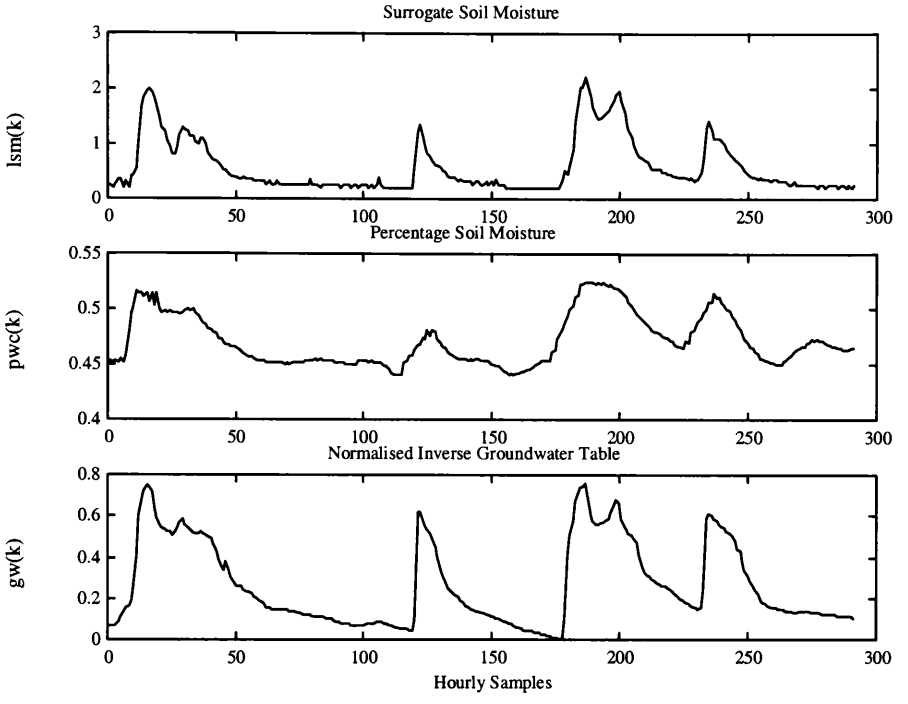

Figure 4.11. (a) Surrogate soil moisture content $lsm(k)$ (b) percentage soil moisture content $pwc(k)$ and (c) normalised inverse depth to groundwater table, Lancaster DBM model $gw(k)$, Data Series 1.

A comparison of $lsm(k)$ versus $pwc(k)$ and $gw(k)$ is presented in Figure 4.12. This straightforward regression suggests the soil moisture variables $pwc(k)$ and $gw(k)$ are directly related to $lsm(k)$ in an *approximately* linear way. Simple least squares (LS) could be adopted at this stage of the analysis to obtain a quantitative estimate of these relationships. However, a relationship estimated in this manner could be significantly biased due to the presence of measurement noise on the data. An alternative and interesting approach is to investigate the nature of the relationships between $lsm(k)$ and both $gw(k)$ and $pwc(k)$ by adopting, once again, a time varying parameter (TVP) approach (see Section 4.2.1).

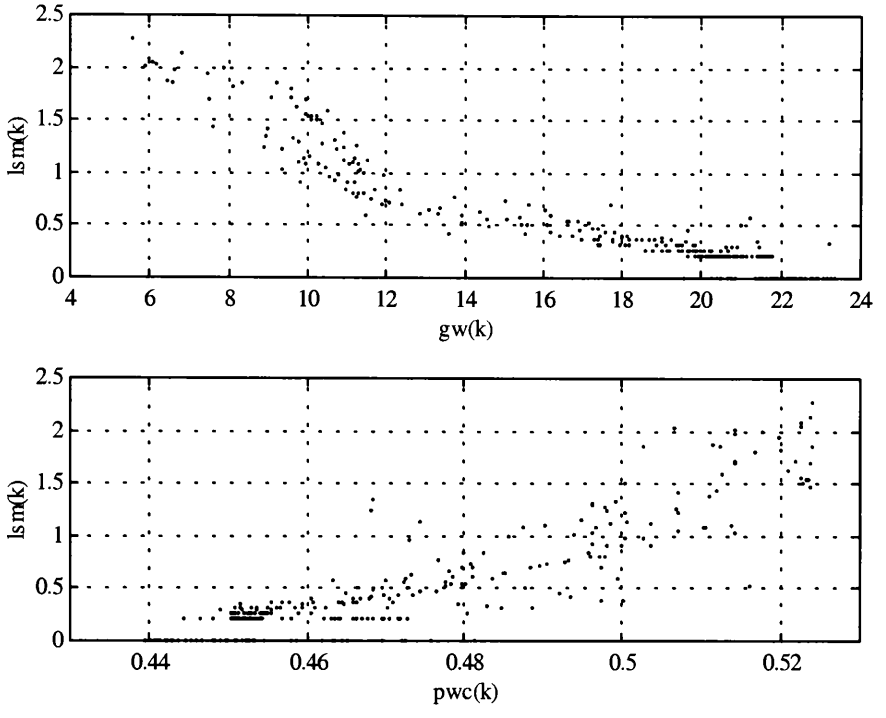

Figure 4.12 Simple regression of $lms(k)$ versus $gw(k)$ and $pwc(k)$

Consider equation 4.18, where $lsm(k)$ is multiplied by a simple gain parameter at each sampling instant such that the product equals $gw(k)$.

$$gw(k) = f(k)\, lsm(k) \tag{4.18}$$

An estimate $\hat{f}(k|N)$ of the TVP $f(k)$ can be obtained over the complete sampling interval from the FIS algorithm and $\hat{f}(k|N) \cdot lsm(k)$, an estimate of $gw(k)$, can subsequently be generated. The nature of the relationship between $\hat{f}(k|N) \cdot lsm(k)$ and $gw(k)$ can now be investigated and a variety of functions estimated using weighted least squares (WLS). The covariance matrix $\hat{P}(k|N)$, generated from the FIS algorithm, is incorporated within WLS and specifies the time intervals over which

the TVP $\hat{f}(k|N)$ is badly defined. The corresponding data points at these time intervals can then be assigned a low weighting. The resulting functions estimated from WLS will, therefore, be based on the data points with a higher weighting and hence strong affinity for one other. Once a function has been found that satisfactorily describes the relationship between $\hat{f}(k|N)\cdot lsm(k)$ and $gw(k)$, it can be tested by applying it directly to $lsm(k)$.

As discussed previously in Section 2.4.2, a noise variance ratio (NVR) must be introduced into the FIS algorithm. If a large NVR is chosen, $\hat{f}(k|N)$ will change rapidly and the product $\hat{f}(k|N)\cdot lsm(k)$ will exactly match $gw(k)$. However, the associated covariance matrix will be badly defined over the complete sampling interval and any function estimated using WLS will be no more discriminatory than simple LS. Moreover, when this function is tested by applying it directly on $lsm(k)$, the resulting time series is likely to give a poor fit to $gw(k)$. An optimal value for the NVR must, therefore, be estimated for use in this analysis and can be obtained from prediction error decomposition (see Section 2.5.2).

Following the procedures described above, an investigation was undertaken to establish the best function describing the relationship between $lsm(k)$ and $gw(k)$, and $lsm(k)$ and $pwc(k)$ respectively. Having estimated $\hat{f}(k|N)$, a scatter plot of $\hat{f}(k|N)\cdot lsm(k)$ versus $gw(k)$ was generated which visually revealed a much closer defined relationship between the two series than is presented in Figure 4.12. From these two series, both a linear function and a variety of nonlinear functions were

estimated. This whole process was repeated using the percentage soil moisture series $pwc(k)$ and a further set of linear and nonlinear functions estimated.

The best linear and nonlinear functions describing the relationship between the soil moisture surrogate $lsm(k)$ and soil water variables $gw(k)$ and $pwc(k)$ are defined in equations (4.16-4.19) and presented in Figures 4.13-4.14.

$$gw(k) = -6.69lsm(k) + 19.68 \tag{4.19}$$

$$gw(k) = -10.80lsm(k)^{-0.49} \tag{4.20}$$

$$pwc(k) = 0.087lsm(k) + 0.399 \tag{4.21}$$

$$pwc(k) = 0.50lsm^{0.17} \tag{4.22}$$

The output from these four functions are plotted (y axis) opposite their respective soil water variable (x axis) on the scatter plots shown in Figure 4.13. Note that the straight line on each plot indicates the ideal solution which the estimated functions are emulating. Taking into account the scale differences, the scatter of data shown in Figures 4.13a and 4.13b has a lower variance and resides closest to the 'perfect line' in comparison with the data presented in Figures 4.13c and 4.13d for $pwc(k)$. Particularly, the nonlinear power function (equation 4.20, Figure 4.13b) generates the best explanation of the relationship between $lsm(k)$ and $gw(k)$, where the scatter of data closely follows the 'perfect line'. In contrast, the linear function (equation 4.19 and Figure 4.13c) gives the best explanation of the relationship between $lsm(k)$ and $pwc(k)$. Overall, the scatter plots show the soil moisture surrogate $lsm(k)$ exhibiting the greatest correlation to $gw(k)$.

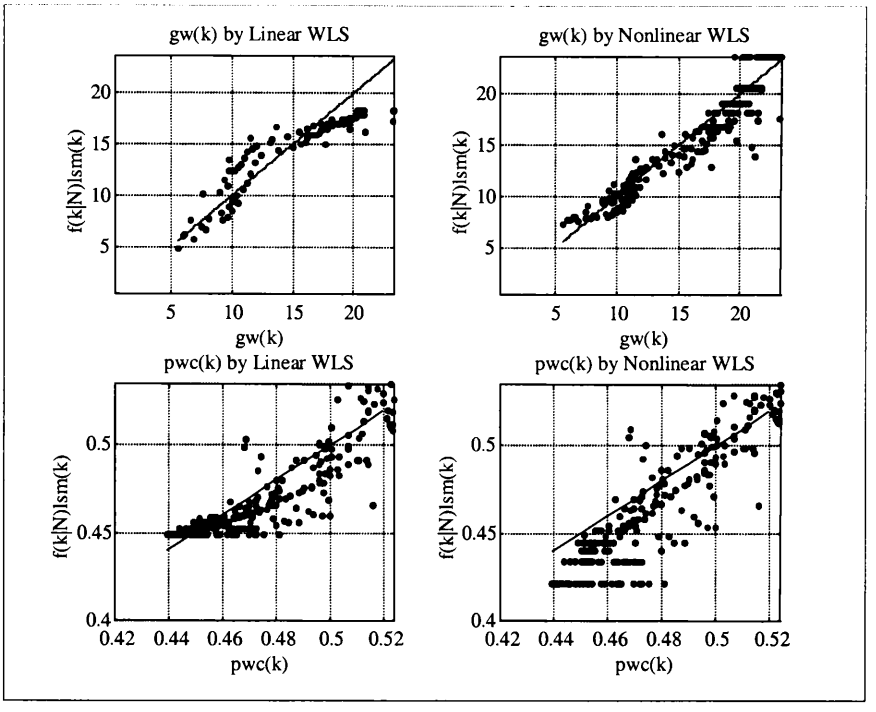

Figure 4.13. Estimated and measured groundwater table depth $gw(k)$ (a: Top Left b: Top Right) and soil moisture content $pwc(k)$ (c: Bottom Left d: Bottom Right) from linear and nonlinear relationships, Lancaster DBM model, Data Series 1.

The ability of each of the four estimated functions (equations 4.19-4.22) to replicate the dynamics of the soil water variables $gw(k)$ and $pwc(k)$ can, perhaps, be more clearly represented in Figure 4.14. Here, the time series generated from each of the four functions are plotted alongside $gw(k)$ and $pwc(k)$ respectively. As expected, the time series generated from the nonlinear function (equation 4.20) shows a much closer fit to $gw(k)$ over the complete range of depths, than the linear function (see Figure 4.14a). Likewise, the dynamics of $pwc(k)$ (see Figure 4.14b) are better replicated by the time series generated from the linear function (equation 4.21).

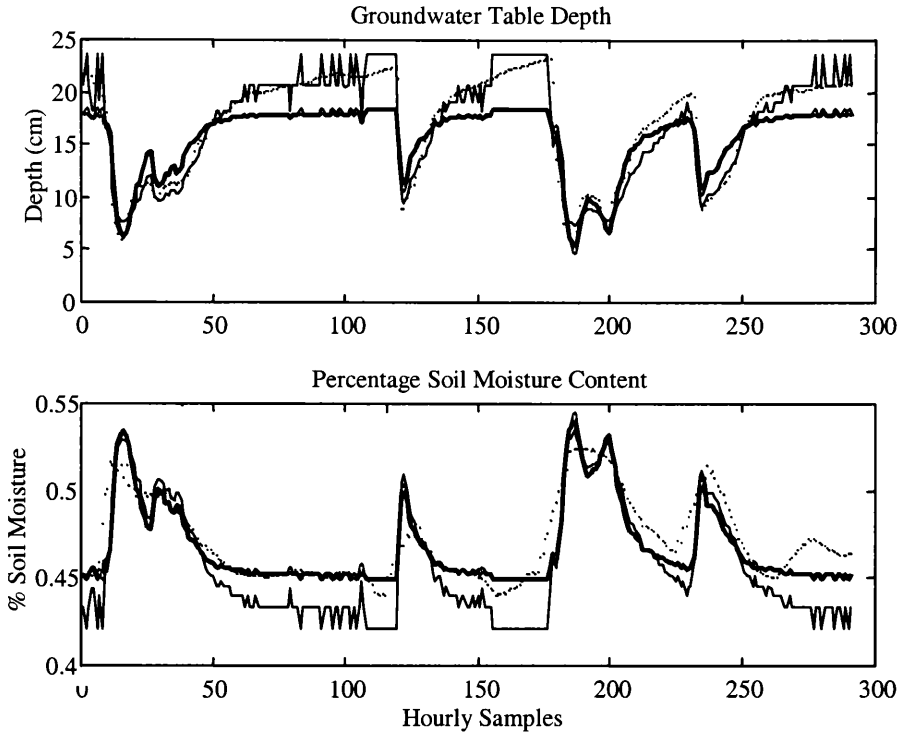

Figure 4.14 Estimates of (a: top) depth to groundwater table and (b: bottom) soil moisture content from linear and nonlinear, WLS relationships, Lancaster DBM model, Data Series 1. Data (Dots); Linear relat. (Bold line); Nonlinear relat. (Fine line).

Figures 4.13 and 4.14 indicate that the time series generated from the four functions show the strongest correlation to the measured soil water variables $gw(k)$ and $pwc(k)$ during wet conditions. Not surprisingly the TVP estimate $\hat{f}(k|N)$ is better defined statistically when more information is contained in the data which, in this case, occurs when rainfall is exciting the system. The precipitation series used to generate $lsm(k)$ includes periods where there is no rainfall. Over these periods, the TVP estimate $\hat{f}(k|N)$ is badly defined, as reflected in the weighting matrix used in the WLS estimation. This accounts for the poorer coincidence between the generated

and measured series during these drier conditions. Since the behaviour of water in the soil system becomes more complex and nonstationary during low flow conditions, it was felt initially, that the relationships could be improved by disregarding these periods of comparative inactivity and exploring the relationship only where the runoff was above a defined threshold. However, this yielded no substantial improvement.

4.4.2 Comparison of the measured soil water variables with the IHACRES soil moisture surrogate

The IHACRES soil moisture surrogate $ihsm(k)$ is shown plotted with the soil water variables $pwc(k)$ and $gw(k)$ in Figure 4.15.

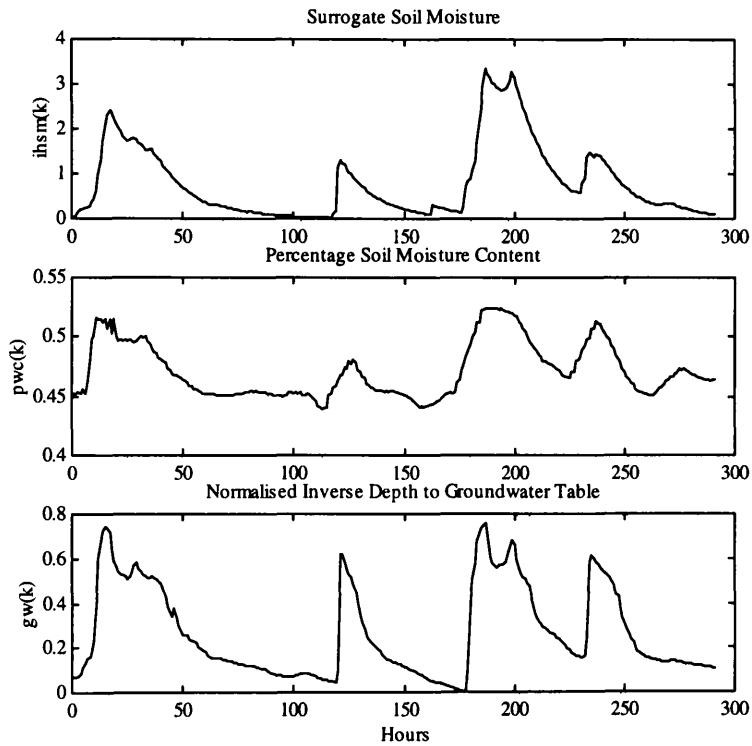

Figure 4.15. (a) Surrogate soil moisture content $ihsm(k)$, (b) percentage soil moisture content $pwc(k)$, (c) normalised inverse depth to groundwater table $gw(k)$, Data Series 1.

The $ihsm(k)$ dynamics show similarity to both soil water variables, but it is evident, once again, that $ihsm(k)$ exhibits greater correlation to $gw(k)$. The soil moisture surrogate $ihsm(k)$ does show evidence of the double peaks which are present within the groundwater table series $gw(k)$. However, the magnitude of the estimated storm peaks do not match the data well. This is particularly evident at the 200 hour time interval, where the $ihsm(k)$ peak well exceeds the remaining series. This characteristic is not present within the $gw(k)$ time series. Note also, in comparison to $gw(k)$ that the recessions of $ihsm(k)$ fall in a much smoother, exponential manner.

Simple scatter plots of $ihsm(k)$ versus the soil water variables $gw(k)$ and $pwc(k)$ do not reveal relationships that are as well defined as for the Lancaster surrogate $lsm(k)$. Having estimated the TVP $\hat{f}(k|N)$, a number of linear and nonlinear functions describing the relationship between $\hat{f}(k|N) \cdot ihsm(k)$ and both $gw(k)$ and $pwc(k)$ were estimated using WLS. The best linear and nonlinear functions describing this relationship are defined in equations (4.23-4.26) and presented in Figure 4.16.

$$gw(k) = -3.36 ihsm(k) + 18.16 \tag{4.23}$$

$$gw(k) = -3.10 ihsm(k)^{-0.27} \tag{4.24}$$

$$pwc(k) = 0.03 ihsm(k) + 0.45 \tag{4.25}$$

$$pwc(k) = 0.46 ihsm(k)^{0.06} \tag{4.26}$$

Figure 4.16 emphasises how each of these four functions are unable to characterise the relationship between $ihsm(k)$ and the soil water variables; importantly, each scatter plot appears to show the data deviating more significantly from the 'ideal line' than is the case of the Lancaster soil moisture surrogate $lsm(k)$ (see Figure 4.13). Overall, the analysis indicates that $ihsm(k)$ shows the greatest correlation to $pwc(k)$ with a linear function explaining this relationship.

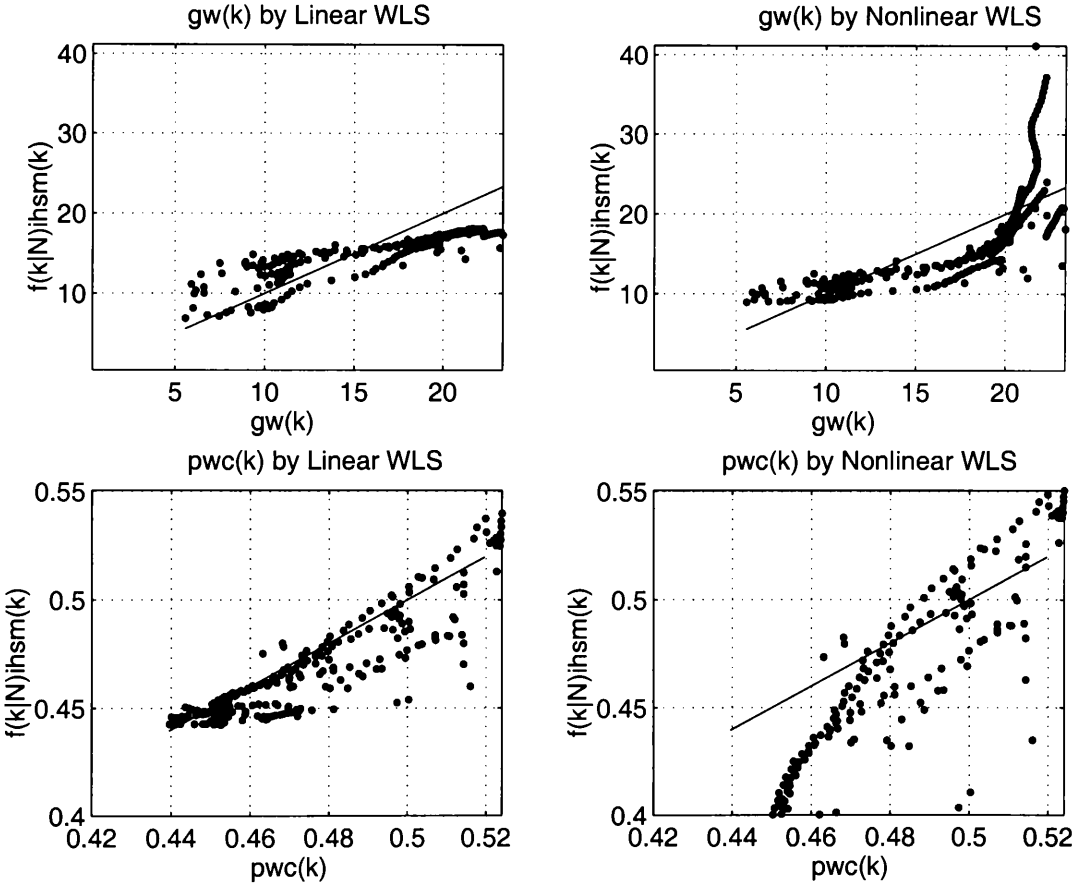

Figure 4.16 Estimate and measured depth to groundwater table and soil moisture content from linear and nonlinear relationships, IHACRES model, Data Series 1.

4.4.3 Data Series 2

Having identified and estimated both rainfall-flow models for Data Series 1 (Section 4.3.3 and 4.3.4) and analysed their respective surrogate soil moisture series (Sections 4.4.1 and 4.4.2), the whole procedure was repeated for Data Series 2.

From an initial investigation it is evident, in the case of the Lancaster DBM model, that the non parametric FIS estimate of $\hat{b}_0(k \mid N)$ can be related to flow by a nonlinear power law. Optimisation of the power law ($\beta = 0.206$) results in the best identified nonlinear model as first order [1,2,0] with $R_T^2 = 0.692$, $YIC = -4.116$, and $AIC = -1.743$. The SRIV algorithm identifies the best nonlinear IHACRES TF model as first order [1,2,0] with model statistics $R_T^2 = 0.779$, $YIC = -5.744$, and $AIC = -2.069$, where the nonlinear module parameters are iteratively optimised to $T_w = 34.513$ and $c = 0.042$. These statistics demonstrate that both models are unable to provide as good a fit to the flow series as was the case with Data Series 1. Noticeably, the IHACRES model gives a superior fit to the flow series than the Lancaster DBM model. This is primarily because the soil moisture surrogate generated by IHACRES model is able to characterise the extremely complex and nonlinear dynamics of the storm event occurring at 300-350 hours (Figure 4.3) in a more successful manner than the Lancaster DBM model.

4.4.4 Comparisons between Data Series 1 and Data Series 2

There is a significant difference in the Lancaster DBM TF (excluding the nonlinear module) between the model parameters estimated from Data Series 1 and those

estimated from Data Series 2; hence, a model estimated from Data Series 1 could not be validated successfully on Data Series 2 and vice versa. Similarly, the parameters of the two IHACRES TF models (including the nonlinear module) have significant disparity. Ideally, one model could be estimated to describe a range of data series. It is possible however, that if a longer time series were available, encompassing a greater range of the catchment antecedent conditions, it would be more likely that a generic model could be estimated.

Comparison of the newly defined soil moisture surrogate $lsm(k)$ generated from Data Series 2 by the Lancaster DBM model, with the groundwater table depth $gw(k)$ shows, once again, the best defined relationship is a nonlinear one. Significantly, this relationship closely resembles that found for Data Series 1, as shown in equation (4.27).

$$\begin{aligned} gw(k) &= -10.80\,lsm(k)^{-0.49} \qquad \text{Data Series 1} \\ gw(k) &= -12.35\,lsm(k)^{-0.56} \qquad \text{Data Series 2} \end{aligned} \tag{4.27}$$

These nonlinear relationships based, as they are, on quite separate data sets, suggest that it may be possible to define a generic nonlinear module for the model if additional time series data can be obtained and analysed.

A simple linear model defines the best relationship between the IHACRES soil moisture surrogate $ihsm(k)$ and $gw(k)$ for Data Series 2, as was found to be the case with Data Series 1. The resulting new series, generated from this linear law, characterises the groundwater dynamics well during wet conditions, however, it does

cause a significant overestimation of the groundwater depth during dry periods. The two linear relationships defined for Data Series 1 and 2, show no real approximation to one another (see equation. 4.28)

$$\begin{aligned} gw(k) &= -3.36 ihsm(k) + 18.16 \quad \text{Data Series 1} \\ gw(k) &= -13.51 ihsm(k) + 28.27 \quad \text{Data Series 2} \end{aligned} \tag{4.28}$$

4.4.5 Summary for nonlinear modelling for Data Series 1 and 2

From analyses of both Data Series 1 and 2 it is clear that the dynamics of the surrogate soil moisture series generated from both the Lancaster DBM and IHACRES models exhibit similar trends to the soil water variables $gw(k)$ and $pwc(k)$. In particular, the Lancaster DBM soil moisture surrogate $lsm(k)$ shows a closer resemblance to $gw(k)$ and $pwc(k)$ than series $ihsm(k)$ generated from the IHACRES model. In both cases, however, the best correlation was achieved with both $gw(k)$ series which may be due, in part, to greater certainty regarding the measurement of groundwater table data. In particular, the interaction of water with soil is dependent on the soil properties: quite large variations of soil moisture content can be expected from point measurements made at plot scale, even within a homogenous soil, because of localised differences in the soil properties. However, spatial variations in ground water level at the same scale are far less significant because of the infinite number of contributing sources smoothing out the localised variations.

It has been assumed so far that the FIS algorithm has been able to correctly identify the complete nonlinearity in the rainfall-flow system from the time series. Under the assumption that $gw(k)$ reflects an accurate representation of the antecedent conditions

of the plot, the fact that the soil moisture surrogate $lsm(k)$ can be related to $gw(k)$ by a nonlinear function, may suggest that the TVP/SDP analysis, undertaken in Section 4.3.3, may not have quite revealed the *complete* rainfall-flow nonlinearity. However, if an enhanced model fit is not obtained from inserting $gw(k)$ directly into the Lancaster DBM model, this might indicate that the soil moisture surrogate $lsm(k)$ does, in fact, reflect a more representative estimate of the complete rainfall-flow nonlinearity (e.g. including water losses to groundwater). This will be investigated in following sections of this chapter.

4.5 FURTHER TVP ANALYSIS

As discussed in Section 3.3.2, more recent nonlinear research has shown that sorting the data prior to the estimation of the TVP can significantly improve the identification of a nonlinearity present in the time series. In light of this development and following the results discussed in the Section 4.3.3, it was considered prudent to repeat the TVP estimation stage of the DBM modelling procedure for the Lancaster model, so as to either reinforce the state dependent relationships already identified or to highlight any new ensuing relationships. All previous applications of this new methodology have sorted the data in terms of the input (see e.g. Section 3.3.2 or Young (1999)). However, in the case of rainfall-flow data, this method of estimation does not help to improve the identification of a state dependent relationship. In contrast, if the data are sorted in terms of the output, a notably enhanced state dependent relationship can be observed.

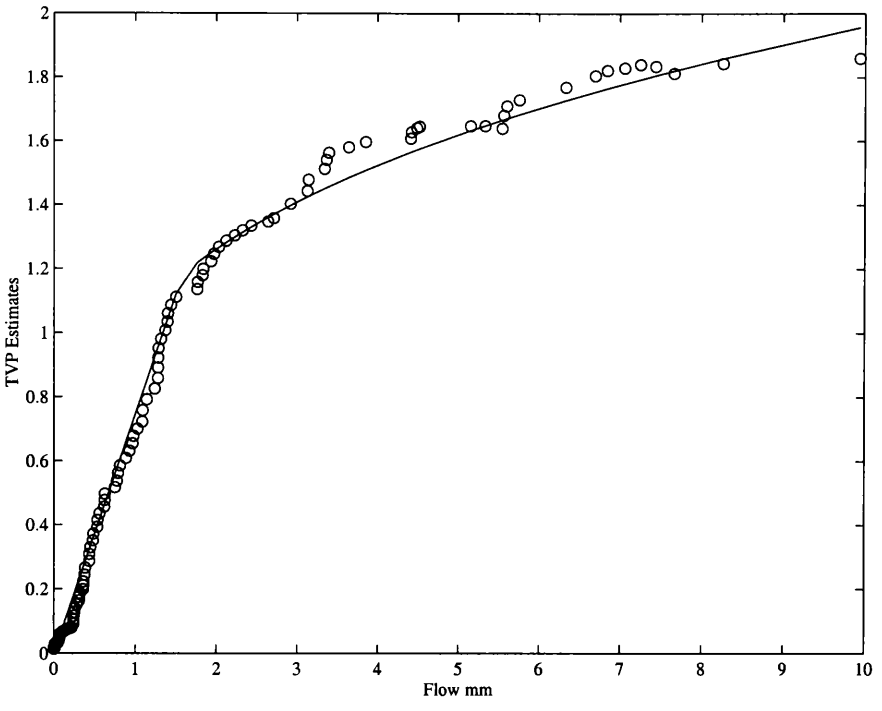

Figure 4.17 The most significant estimates of $\hat{b}_0(\circ \mid N)$ versus flow and the WLS estimate of two stage function comprising of a linear and nonlinear relationship, Lancaster DBM model, Data Series 1

The relationship identified for Data Series 1 is presented in Figures 4.17 which strongly suggests that a two stage function (equation 4.29) may describe the relationship between the TVP $\hat{b}_0(\circ \mid N)$ and $y(k)$.

$$\begin{aligned} \hat{b}_0(\circ \mid N) &= 0.747 y(k) \quad \text{for } y(k) < 1.51 \\ \hat{b}_0(\circ \mid N) &= 0.045 + y(k)^{0.282} \quad \text{for } y(k) > 1.51 \end{aligned} \quad \text{Data Series 1} \qquad (4.29)$$

Although a two stage relationship of this kind has previously been identified from rainfall-flow time series (Young, 1993), it is only through sorting the data, that it has revealed such a well defined state dependent relationship. As a result, only the

statistically well estimated TVP's determine the shape of the state dependent relationship.

In physical terms it is considered that the first component of the relationship determines the flow response when the catchment is extremely dry. Rainfall will be readily taken up by the soil to reduce the large soil moisture deficit and as a result has a minimal affect on the stream flow response. With continued rainfall, the soil moisture deficit rapidly decreases and the flow response to rainfall increases in a linear manner up to a threshold, determined by the specific physical properties of the soil. Beyond this threshold (the second component of the relationship), when the soil moisture status of the soil is higher, the deficit is reduced much more gradually until the soil reaches saturation.

A more clearly defined state dependent relationship (see equation 4.30) is also identified by re-estimating the TVP for Data Series 2 utilising sorted data.

$$\hat{b}_0(\circ \mid N) = 0.293 y(k)^{0.276} \quad \text{Data Series 2} \tag{4.30}$$

However, on this occasion, this approach proves to clarify further the state dependent relationship identified in Section 4.4.3, rather than revealing a new alternative relationship, as determined with Data Series 1.

4.5.1 Model fits with sorted data

Although the re-evaluated nonlinear state dependent relationships are more clearly defined, they do not drastically enhance the model fit. For example, utilising the two stage relationship rather than the single power law for Data Series 1, a model with only a slightly superior fit is identified with the following statistics: $R_T^2 = 0.974, YIC = -6.124$, and $AIC = -2.615$. Similarly, this re-evaluated nonlinear relationship does not improve the correlation between the resulting newly defined soil moisture surrogate $lsm(k)$ (see equation 4.31) and the soil water variables $gw(k)$ and $pwc(k)$.

$$\begin{aligned} lsm(k) &= 0.747 y(k) && \text{for } y(k) < 1.51 \\ lsm(k) &= 0.045 + y(k)^{0.282} && \text{for } y(k) > 1.51 \end{aligned} \tag{4.31}$$

An initial visual comparison of $lsm(k)$ and the soil water variables $gw(k)$ and $pwc(k)$, shown in Figure 4.18, indicates that the level of correlation existing between these series is not, as originally expected, higher than that found from the previous analysis (Section 4.4.1). The shape of the recessions show the most dissimilarity; $lsm(k)$ has much shorter recessions than both $gw(k)$ and $pwc(k)$. Note also from Figure 4.18 that the recessions of $lsm(k)$ are much shorter than the soil moisture surrogate series previously defined in Section 4.4.1. These shorter recessions are caused by the first component of the two stage function (equation 4.31) reducing the value $lsm(k)$ far more than the power law defined by equation (4.16).

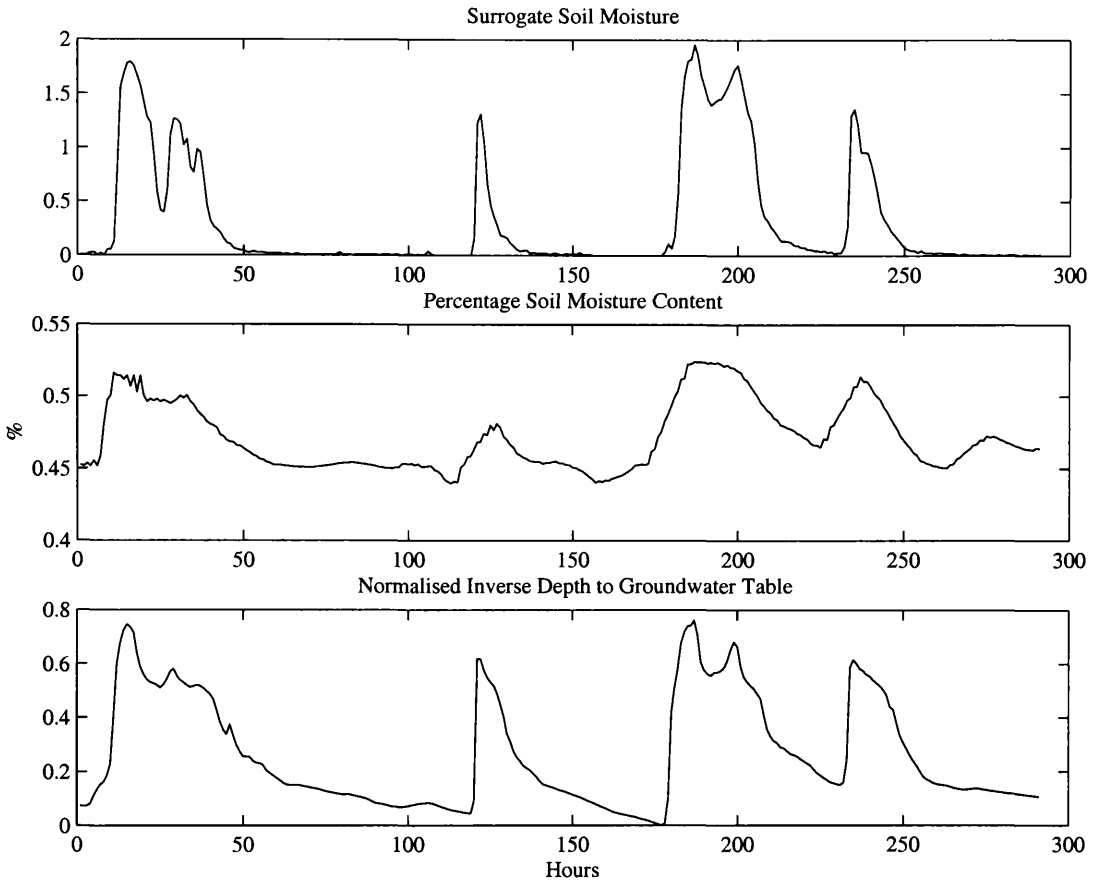

Figure 4.18 (a) Soil moisture surrogate $lsm(k)$, (b) percentage soil moisture content $pwc(k)$ and (c) normalised inverse depth to groundwater table $gw(k)$, Lancaster DBM model, Data Series 1.

4.6 LANCASTER DBM MODEL MODIFICATION

As previously discussed, hydrological monitoring programs rarely collect detailed soil water time-series. Consequently, the Swiss plot-scale data offers an excellent opportunity to establish whether the soil water variables $gw(k)$ and $pwc(k)$ can be incorporated directly into the Lancaster DBM rainfall-flow model to replace the soil moisture surrogate. It was demonstrated in Sections 4.3.1 and 4.4.3 that the rainfall-flow TF model parameter $\hat{b}_0(k|N)$ exhibits significant temporal variation. On this

occasion, rather than relating the temporal variation of $\hat{b}_0(k|N)$ to $y(k)$ (c.f. Young, 1993), the TVP is compared directly to the soil water variables $gw(k)$ and $pwc(k)$. Well defined nonlinear relationships are identified between $\hat{b}_0(k|N)$ and both soil water variables from Data Series 1, but only $gw(k)$ from Data Series 2. A typical example is given in Figure 4.19 for Data Series 1, which shows the most significant estimates of $\hat{b}_0(k|N)$ plotted against $pwc(k)$ with the WLS estimate of this relationship.

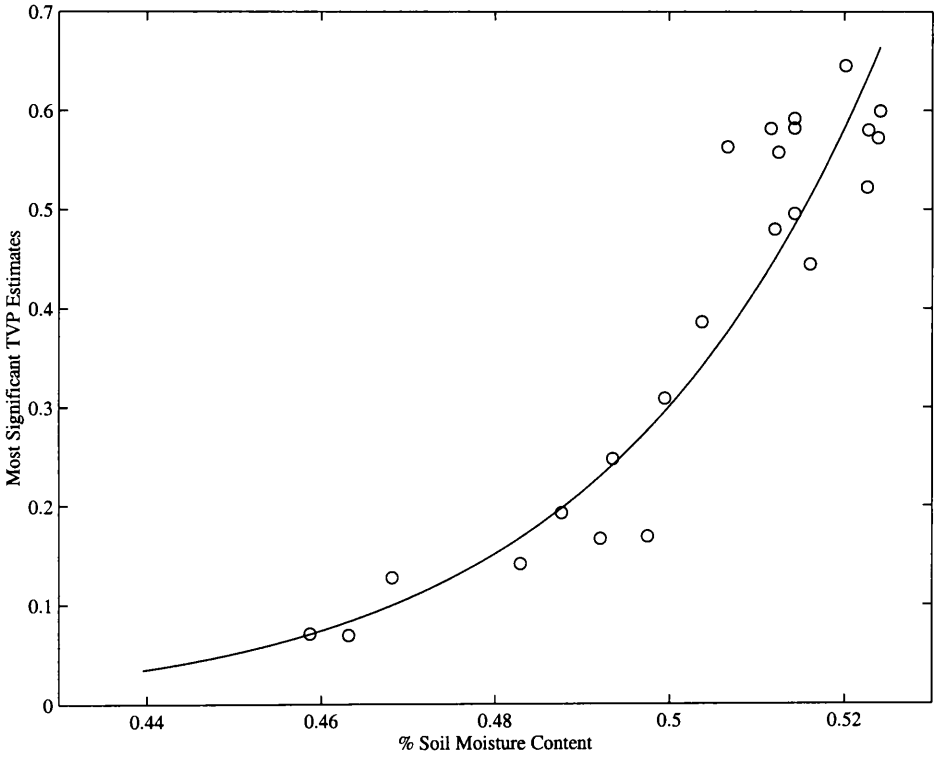

Figure 4.19 The most significant estimates of $\hat{b}_0(k \mid N)$ versus $pwc(k)$ and the WLS estimate of the power law relationship.

Having identified the nature of the nonlinearity and hence the structure of the nonlinear model, statistically efficient estimates of the model parameters are obtained through a process of iterative optimisation incorporating the SRIV algorithm. Tables

4.3 and 4.4 present the statistical diagnostic criteria for the best nonlinear models estimated for Data Series 1 and 2 respectively.

Table 4.3. Nonlinear Lancaster DBM models, Data Series 1.

Nonlin.	Structure	Power Law	*YIC*	R_T^2	*AIC*
$y(k)$	[1,3,0]	0.358 (0.007)	-6.822	0.966	-2.337
$gw(k)$	[1,3,0]	-1.146 (0.069)	-6.119	0.951	-1.977
$pwc(k)$	[1,2,1]	12.061 (0.739)	-6.142	0.948	-1.923

Table 4.4. Nonlinear Lancaster DBM models, Data Series 2.

Nonlin.	Structure	Power Law	*YIC*	R_T^2	*AIC*
$y(k)$	[1,2,0]	0.206 (0.012)	-4.116	0.692	-1.743
$gw(k)$	[1,3,0]	-1.898 (0.075)	-5.948	0.863	-2.546

Table 4.3 highlights that the two models incorporating the soil water variables provide a very good fit to Data Series 1, although the nonlinear model incorporating $gw(k)$ gives a slightly better fit. These initial results indicate, that for modelling purposes, the generated soil moisture surrogate $(y(k)^{0.358})$ provides, in this instance, a better indication of the overall antecedent conditions of the plot than the actual soil moisture measurements $gw(k)$ and $pwc(k)$. However, the modelling results from Data Series 2, provide a contrasting view (see Table 4.4), as on this occasion, the inclusion of $gw(k)$ directly into the model, significantly enhances the model fit (see Figure 4.20). There is a need for more research into this alternative modelling approach, but the results do indicate that when available, the inclusion of the soil water variables directly into the model could enhance the model performance and enhance the physical interpretation of the model.

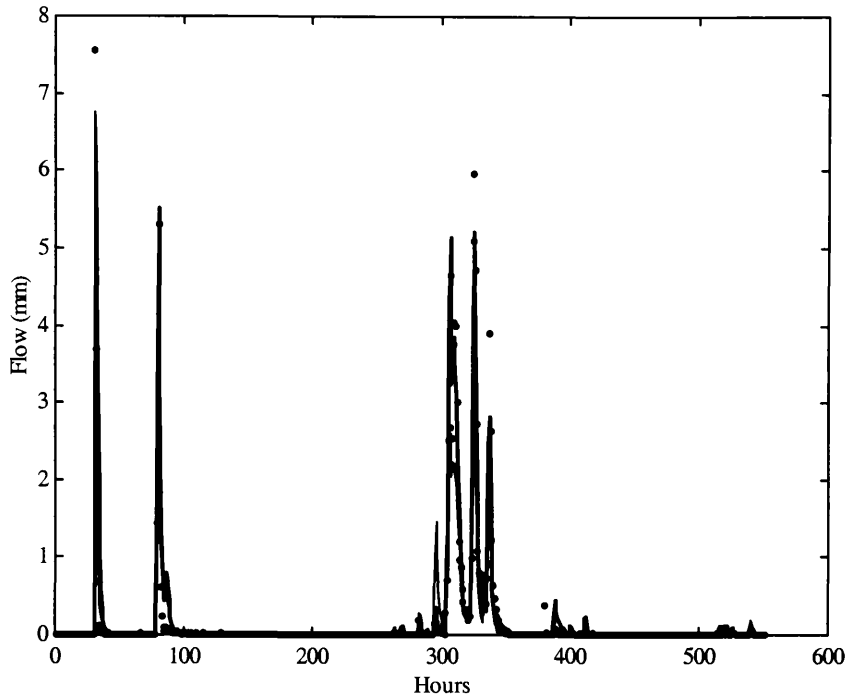

Figure 4.20 Best nonlinear Lancaster DBM TF models , Data Series 2.

Model utilising $gw(k)$ (bold line), utilising $y(k)$ (fine line), and measured data (dots).

4.7 CONCLUSION

This chapter has critically evaluated how efficiently the surrogate for soil moisture conditions, in both the IHACRES and Lancaster DBM rainfall-flow models, captures the actual antecedent soil moisture conditions as defined by the field data collected at the plot scale in the Swiss pre-Alps.

The theoretical background to both models has been described and models have been identified and estimated from two independent data series. The surrogate soil moisture series generated by both models capture the dominant dynamics of the measured soil moisture variables. However, overall, the soil moisture surrogate

generated by the Lancaster DBM model (Young, 1993) exhibits the highest correlation to the measured soil moisture variables for both Data Series 1 and Data Series 2, particularly to the groundwater table time series (see also Fawcett *et al.*, 1997).

This analysis validates the assumptions made by the Lancaster DBM model that flow can be used to define an efficient surrogate for catchment antecedent conditions with the resulting bilinear TF models explaining the rainfall-flow behaviour very well. As a result, it is considered that the DBM rainfall-flow model, which relies only on rainfall and flow records, can be applied with confidence to catchments where the more extensive soil moisture and depth to groundwater table time series are unavailable. However, the research also suggests that when these additional measures are available, the direct use of the ground water depth series may improve the model performance and its physical interpretation.

An improved method for identifying the nonlinearity in rainfall-flow series has also been introduced. A much better defined state dependent relationship between flow and $\hat{b}_0(k \mid N)$ is identified if the data are sorted in ascending order with respect to flow rather than rainfall, before estimating the TVP.

Conclusions drawn from the research at the plot scale can only be cautiously related to catchment scale processes. The Swiss data collected at the plot scale offers the first opportunity to critically evaluate the soil moisture surrogates generated from the two rainfall-flow models. Positive conclusions drawn from this research should act as a further impetus to consider work investigating the issues of upscaling.

CHAPTER 5

MODELLING FLOW, SUSPENDED SEDIMENT LOAD AND RESERVOIR SEDIMENTATION USING A DATA-BASED MECHANISTIC (DBM) APPROACH; WYRESDALE PARK CATCHMENT, LANCASHIRE, UK

Suspended sediment has a highly detrimental impact on hydraulic structures and ecological systems and as a result, there is a basic requirement to be able to accurately describe and predict sediment load at key sites within a catchment. In particular, suspended sediment is an important consideration in the design life and operation of reservoirs as it affects both water quality and storage capability. It is therefore necessary, where possible, to generate accurate predictions of suspended sediment and understand the relationship between catchment practises and sediment yield. The previous chapter has introduced the data-based mechanistic (DBM) methodology and

how it can be utilised to effectively identify and estimate the nonlinearity in the hydrological rainfall-flow model. This chapter discusses a novel application of the objective DBM approach to identify a nonlinear, low-order lumped-parameter model relating rainfall to stream suspended sediment load (SSL) at Wyresdale Park reservoir inflow with the overall aim of modelling historical reservoir sedimentation.

Sediment delivery through the catchment comprises a mesh of complex nonlinear processes. It is therefore difficult, where data is in short supply, to develop a physically based or highly mechanistic model which requires extensive field work to validate. Reservoir sedimentation provides an important record of sediment yield. At Wyresdale there exist data sets of rainfall, sediment concentration and discharge time series and bathymetric surveys of the lake. Such data is ideal for a DBM study and unlike other reservoir based studies of historical reservoir sedimentation, (e.g. Labadz *et al.*, 1991; Curr, 1985; Foster *et al.*, 1986) the DBM approach can provide temporal data on annual or seasonal scales. In contrast to physically based models and previous reservoir studies, the DBM modelling methodology discussed in this chapter is well suited to complex systems where the input-output behaviour is of prime importance. This approach naturally results in parametrically efficient, robust, low-order models which are inherently stochastic, and so can account for model uncertainty, sampling and measurement errors.

In order to model the SSL series at Wyresdale, a model must address the key nonlinearities of the system. Flow fundamentally affects the transport of sediment within the catchment and it is vital, to be able to characterise the nonlinear stream response to rainfall, which is heavily dependent on the antecedent conditions of the

catchment. Further, there is an additional nonlinear relationship between flow and SSL whereby, flow magnitude and duration governs the capacity of the flow to convey sediment of a particular calibre to the basin outlet. The previous chapter has discussed the Lancaster DBM rainfall-flow model which will be applied here to evaluate and characterise the rainfall-flow dynamics of the Wyresdale catchment. Working from this initial analysis, a rainfall-sediment model is developed where this nonlinearity is characterised in a similar manner to the rainfall-flow model of Young (1993). The rainfall-sediment model forms the key component of a DBM modelling study of historical reservoir sedimentation at Wyresdale Park.

The overall Wyresdale model, comprises the rainfall-sediment model coupled to a nonlinear model, which relates the SSL at the reservoir inflow to SSL at the reservoir spillway. The combined model, is then simulated, using a historical rainfall-series as an input, to reconstruct daily deposition rates at the reservoir between 1911-1996. This synthetic sediment accretion sequence is compared with the variations in sand content within sediment cores collected from the reservoir.

5.1 WYRESDALE PARK CATCHMENT AND DATA.

Wyresdale Park catchment is situated at the edge of the Forest of Bowland, 10km south of Lancaster, Lancashire (SD512294) and covers an area of 3 km^2. Situated within the catchment is Wyresdale Park Reservoir (Plate 5.1) which is shallow and unregulated, with a surface area of only 0.08 km^2 and mean depth of *c.*1.5 m. The reservoir was constructed in 1895 and as shown in Figure 5.1 is fed by Tythe Barn Brook, the principle stream that drains Wyresdale catchment. A small tarn, situated

on Nicky Nook at the head of the catchment (215 m aod) feeds Tythe Barn Brook, which is joined by four minor tributaries before entering the reservoir. As the catchment has remained relatively undisturbed by land management practises it has been the centre of ongoing research projects investigating the temporal and spatial patterns of the supply and deposition of sediment in reservoired catchments (Goodwill, 1998). Tythe Barn Brook was instrumented with continuously logging turbidity meters and stage recorders at the inflow and outflow of the reservoir producing a time series of measurements taken at 15 minute intervals over a two year monitoring period, 1st March 1994 - 28th February 1996 (Goodwill, 1998). Following calibration, this data can be utilised to determine discharge, suspended sediment concentration and suspended sediment load. Further, a continuously logging tipping bucket rain gauge is situated at the head of the catchment on Nicky Nook Fell. A bathymetric survey conducted in 1994 estimated that the storage capacity of the reservoir has been depleted from its original 148 000 m^3 to an estimated 126 000 m^3 with non-uniform sediment deposition patterns (c.f. Goodwill *et al.*, 1995).

Plate 5.1 Wyresdale Park Reservoir, looking north west from the dam wall.

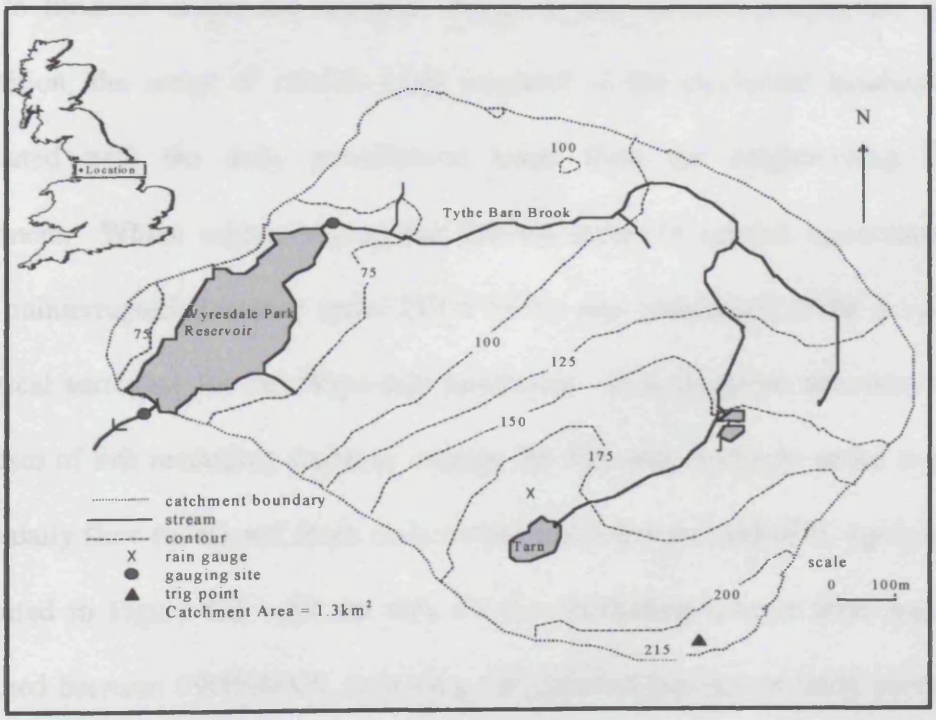

Figure 5.1 Location map of Wyresdale catchment and reservoir

Wyresdale Park catchment was initially considered approximately watertight, with relatively small losses through groundwater to adjacent catchments and moderate evapotranspiration, concurrent with other upland catchments. However, the total rainfall to total flow ratio, calculated from the two years of data, indicate otherwise. The ratios for the two years are 0.17 (1994) and 0.09 (1995), indicating that 83% and 91% of the total catchment precipitation input is lost to either groundwater and/or by evaporation and transpiration. These latter figures were initially thought to be particularly high. However, detailed additional checks utilising the meteorological office rainfall and evaporation calculating system (MORECS) and rainfall and flow data, by Goodwill (1998), supported these ratios.

Historical rainfall records are not available at Wyresdale. Therefore, in order to be able to hindcast suspended sediment transport and reconstruct reservoir sediment deposition, the series of rainfall totals recorded in the catchment headwater were correlated with the daily precipitation totals from the neighbouring Barnacre catchment. Whilst acknowledging the inherent errors of rainfall measurement, the long uninterrupted Barnacre series (1911-1996) was considered to be a very good historical surrogate for the Wyresdale catchment. It is therefore necessary, for the purposes of this modelling study, to average the SSL and discharge series to give the mean daily flow (m^3/s) and mean daily suspended sediment load (SSL kg/s), which is presented in Figure 5.2. All the data for this modelling scheme were mean daily averaged between 0900:0900h, following the standard practice of daily precipitation totals.

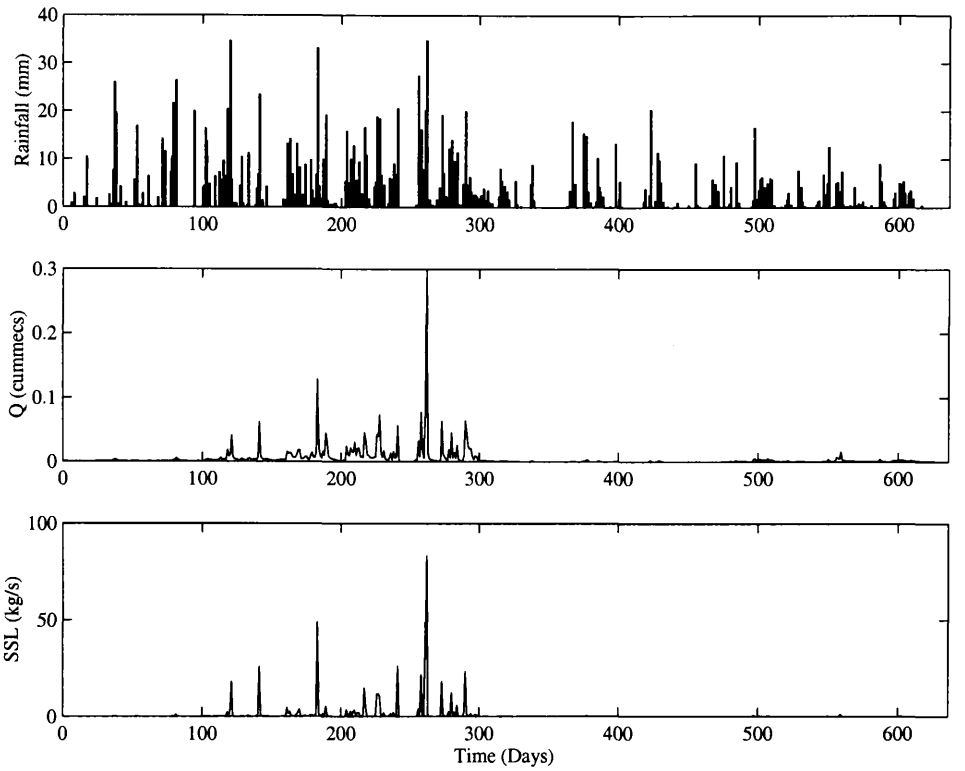

Figure 5.2 Wyresdale Park catchment data. Rainfall measured at Nicky Nook, flow and suspended sediment load measured at the Wyresdale reservoir inflow, 15th May 1994 – 1st February 1996.

5.2 RAINFALL- FLOW MODEL

One of the limitations of developing a hindcasting model for Wyresdale is that the model form is necessarily restricted by the availability of the rainfall data, which is only measured on a daily basis. It follows that any investigation of the nonlinearities in the system must be based on daily averaged data. A further restriction on the model development, is that only two years of data are available, where, obviously for hindcasting purposes, the greatest length of data possible will minimise uncertainties in the model structure. Moreover, the two years of Wyresdale data are extremes, in the sense that 1995 is one of the driest recorded in 40 years at the nearby Hazelrigg

weather station and 1994 is much wetter than average. The majority of large events occur during September-December with 60% of the flow occurring in 7% of the time. Visual inspection of the time series shown in Figure 5.2 supports this information showing predominant seasonal variability in the sediment delivery. Since the two years of data encompass a great range of rainfall and flow conditions, one way of developing a rainfall-flow model is to use the TVP analysis to investigate any underlying seasonal trends in the data using the parameters of the estimated TF models.

5.2.1 Rainfall-flow models on a seasonal scale

In order to look for any seasonal behaviour, the two year record was divided into seven sections according to a 'season' of approximately 90 days in length. The DBM method of time varying parameter (TVP) estimation and state dependent parameter modelling (SDPM), was utilised to identify and characterise the key nonlinearity existing between the rainfall and flow series for each seasonal period of the Wyresdale data set. In each case, the identified nonlinearity can be effectively represented utilising flow within a power law relationship (acting as a surrogate for the antecedent soil moisture conditions of the catchment). The general rainfall-flow model is identified as having the following structure,

$$u_e(k) = u(k)y(k)^{\beta} \tag{5.1}$$

$$y(k) = \frac{B(z^{-1})}{A(z^{-1})}u_e(k-\delta) + \zeta(k) \tag{5.2}$$

where equation (5.1) is the effective rainfall term required to characterise the nonlinearity in the rainfall-flow system. Here, $u(k)$ is daily rainfall in mm; $u_e(k-\delta)$ is the effective rainfall term with pure time delay δ; β is a constant coefficient; $y(k)$ is the mean daily flow (m^3sec^{-1}) at the reservoir inflow; $\zeta(k)$ is a general noise term included to account for stochastic disturbances and unmeasured inputs to the system. The polynomials $A(z^{-1})$ and $B(z^{-1})$ are defined as,

$$A(z^{-1}) = 1 + a_1 z^{-1} + \ldots + a_n z^{-n}; \quad B(z^{-1}) = b_0 + b_1 z^{-1} + \ldots + b_m z^{-m} \tag{5.3}$$

where z^{-1} is the backward shift operator and the integers n and m are the number of parameters in the polynomials.

Table 5.1 presents the best models identified for each seasonal period, and shows that individual models achieve a good fit to their respective calibration data. The effect of daily averaging, combined with the known rapid response of the Wyresdale catchment, reduces the dynamic content of the data which is reflected in the first order model structure identified for each time period.

Table 5.1 Best identified nonlinear rainfall-flow models and the statistical fitting criterion for each seasonal period.

Period	Model Order	Power Law Coefficient	R_T^2	*YIC*	*AIC*	Total Rainfall (mm)
Summer 1994 1/6/94-31/8/95	[1,1,0]	0.68	0.718	-6.524	-15.274	270.9
Autumn 1994 1/10/94-30/11/94	[1,1,0]	0.69	0.910	-6.695	-10.672	356.8
Winter 1994 1/12/94-28/2/95	[1,1,0]	0.65	0.958	-6.694	-10.390	467.6
Spring 1995 1/3/95-31/5/95	[1,1,0]	1	0.920	-6.081	-12.699	156.4
Summer 1995 1/6/95-31/8/95	[1,1,0]	0.6	0.819	-5.124	-17.367	125.2
Autumn 1995 1/10/95-30/11/95	[1,1,0]	1.0	0.785	-6.411	-13.837	163.6
Winter 1995 1/12/95-29/1/95	[1,2,0]	1.00	0.728	-5.378	-15.560	60.40

The identified models were examined to determine whether there are any relationships between any of the estimated model parameters, associated physical descriptors and rainfall totals over the seasonal periods. The pure time delay δ corresponds to the overall lag time between an input of rainfall and the catchment's response. In each model case δ has a value of zero indicating a response time of 0-1 days. This value is small and indicates that the catchment responds very quickly to precipitation, a consequence primarily due to the catchments small areal size, steep slopes and thin soils. The wetter periods of Summer, Autumn and Winter 1994 have lower power function values (~0.67); whilst with the exception of Summer 1995, the unusually dry periods of Spring, Autumn and Winter 1995 have power function values of 1.

Consequently, it could be suggested that the power function values show some time varying behaviour at the seasonal scale, which could be incorporated into a single rainfall-flow model with time varying parameters. Further, a general trend in the values of the denominator (and therefore time constant) and numerator coefficients is apparent; increasing in value, the wetter the period. Whilst these relationships have been identified, additional time series data (e.g. at least 10 years) would be required to be more certain of this time variant behaviour, such that data availability is the limit on further model development.

In a study by Jakeman *et al.,* (1993), TF models for consecutive annual sequences were cross validated in order to evaluate whether one generic model was capable of adequately simulating the response of the catchment over many years. Following a similar procedure with the seasonal TF models, the results shown in Table 5.2 indicate that no one model, estimated from a single seasonal period, is able to adequately simulate the flow characteristic of the catchment over the complete two years of record. The R_T^2 of the models are highlighted along the leading diagonals of the table and the simulation results from passing data from other periods through the models are reported on the off diagonals. For example, when the model calibrated from Autumn 1994 data ($R_T^2 = 0.911$) is simulated using rainfall data from Winter 1995, the simulated flow series explains 0.689% ($R_T^2 = 0.689$) of the observed flow series.

Table 5.2 The performance of calibrated models (horizontal) simulated using data from other seasonal periods (vertical). For example, when the model, calibrated on the data of Summer 1995 ($R_T^2 = 0.819$), is simulated using the Winter 1995 rainfall data, the subsequent simulated flow series compares to the observed flow series with an $R_T^2 = 0.483$.

	Summer 1994 M	**Autumn 1994 M**	**Winter 1994 M**	**Spring 1995 M**	**Summer 1995 M**	**Autumn 1995 M**	**Winter 1995 M**
Summer 1994	0.716	-1.121	-4.036	-3.100	0.561	-3.322	-7.894
Autumn 1994	0.480	0.911	0.814	-6.155	0.604	-6.553	-16.761
Winter 1994	0.390	0.929	0.976	-9.008	0.518	-9.442	-21.703
Spring 1995	0.208	0.412	0.467	0.796	0.198	0.811	0.799
Summer 1995	0.573	0.297	-1.233	-0.010	0.819	-0.135	-2.078
Autumn 1995	0.392	0.637	0.594	0.781	0.403	0.785	0.481
Winter 1995	0.461	0.689	0.647	0.752	0.483	0.760	0.728

The only models that maintain a positive R_T^2 year round are calibrated from Summer 1994 and Summer 1995 data. However, both models completely underestimate the discharge during the wetter periods of Autumn 1994, Winter 1994 and Spring 1995. A seasonal component in the rainfall-flow model structure would be an ideal solution for characterising different seasonal behaviours observed in the time series. However, in view of the diverse climatic conditions reflected in the 2 years of data and the limited temporal extent of this time series, the uncertainties associated with the inference of any seasonal relationships would be too great to warrant adding this extra

complexity into the model. This would suggest, therefore, that the best way to ensure that a reliable model is obtained for the catchment, is to identify a nonlinear TF model based upon the complete 2 year data series.

5.2.2 Rainfall-Flow Model for the 2 year time-series

In order to obtain a single rainfall-flow model for the whole two year time series, the TVP/SDPM procedure was applied once again to identify the key nonlinearity between the rainfall and flow data. Having confirmed that the identified nonlinearity can be characterised by a power function, an optimisation routine incorporating the SRIV algorithm identifies the best model to be first order ($YIC = -7.974$, $R_T^2 = 0.944$, $AIC = -11.172$) and estimates the model parameters to be $a_1 = -0.122\,(9.46e^{-3})$, $b_0 = 0.0189\,(6.32e^{-5})$ and $\beta = 0.46(0.001)$ respectively, where the parameter standard errors are given in parentheses. The model fits the data during both storm peaks and low flows over the majority of the two year series, a section of which is presented in Figure 5.3.

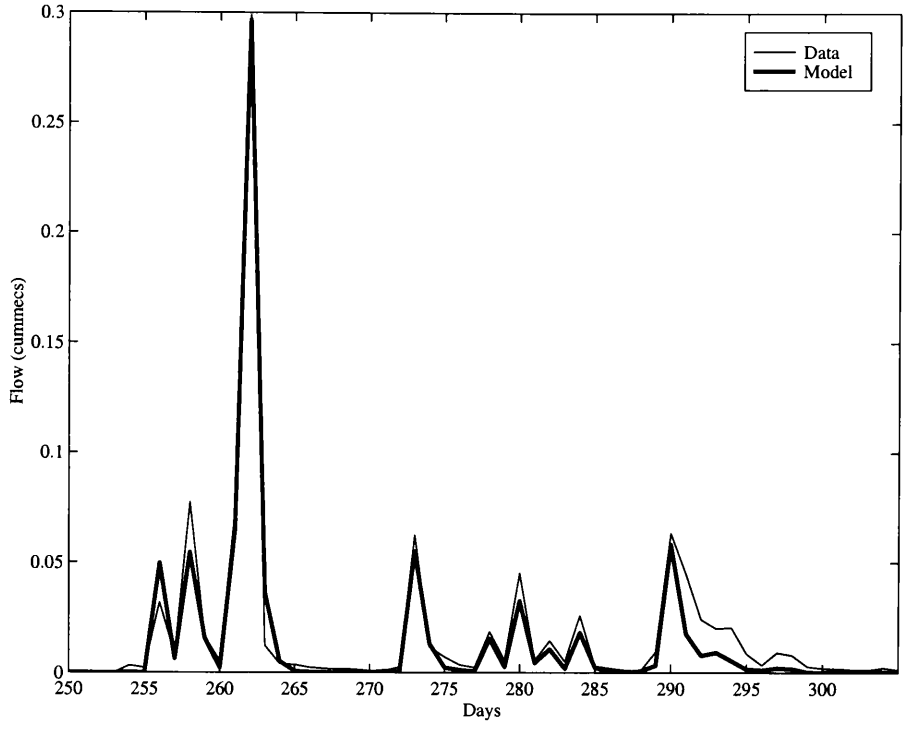

Figure 5.3 A section of the observed two year flow series (fine line) and the nonlinear DBM model. (full line).

5.2.3 Rainfall-Flow Modelling and Evapotranspiration

The effects of evapotranspiration on the rainfall-flow system have so far been regarded as inconsequential at Wyresdale and have been implicitly accounted for within the structure of the model by the nonlinear effective rainfall component. The water balance derivation clearly indicates, however, that evapotranspiration is significant along with groundwater losses at Wyresdale, suggesting that evapotranspiration may need to be explicitly accounted for within the model.

Research from the Coweeta catchment, USA, by Young and Beven (1994) and Young *et al.*, (1997) demonstrated that the residuals of the rainfall-flow model show high correlation with the mean daily temperature series, which in turn may be used as an indicator of actual evapotranspiration. Consequently, an improved model fit can be achieved by developing a multi-input-single-output (MISO) form of the rainfall-flow model, where the mean daily temperature series is introduced as an input to an additional linear TF.

In a similar manner, the residuals of the estimated rainfall-flow model were examined and are presented in Figure 5.4a. The daily temperature series, measured at Lancaster University Hazelrigg weather station (7 km from Wyresdale), corresponding to the two year time period of measured data are presented in Figure 5.4b. Visual inspection of both series indicates that no obvious correlation is evident between them, which has been confirmed from numerical correlation analysis. This indicates that further development of the Wyresdale model to include a temperature component is not required; clearly in this temperate upland catchment the temperature forcing on evapotranspiration is not sufficiently strong to exert a strong influence on the overall catchment dynamic. Whilst this prevents the model from being developed in this regard, the data have, in fact, been thoroughly and objectively analysed and the information utilised to its practical limits to produce the best rainfall-flow model from the available but rather limited data.

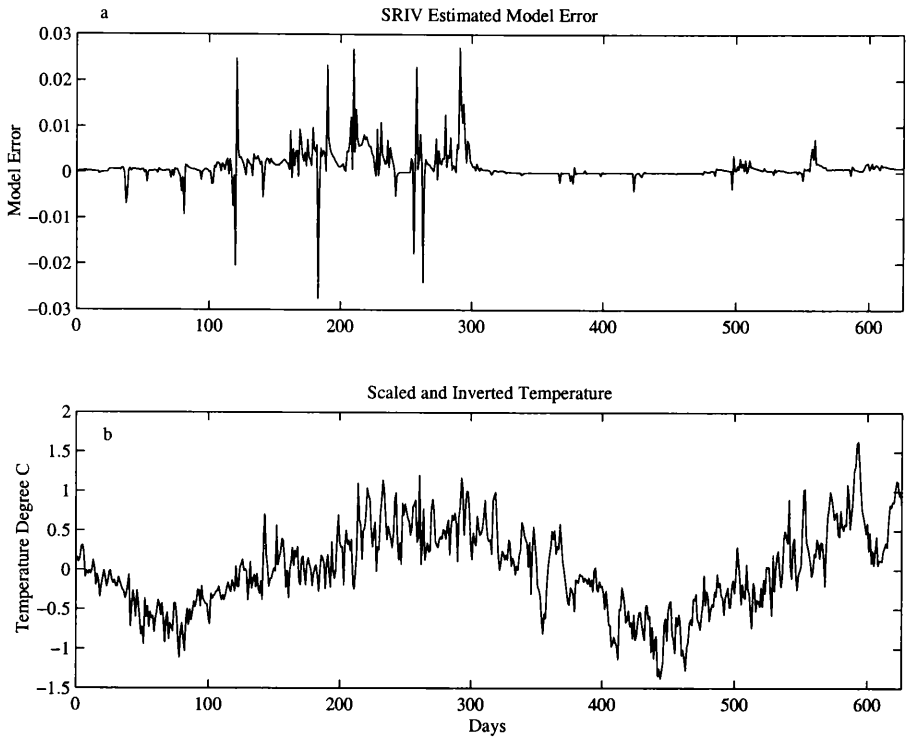

Figure 5.4 (a) SRIV estimated model residuals 15/05/1994 – 01/02/1996, (b) Daily temperature (instantaneous measurement at 0900h), measured at the Lancaster University Hazelrigg weather station for the corresponding time period.

5.3 SUSPENDED SEDIMENT LOAD MODELLING

The key component underpinning the DBM modelling study of reservoir sedimentation at Wyresdale, is the identification of a model relating rainfall to stream SSL at the reservoir inflow. Whilst a great deal of research has been directed at modelling the rainfall-flow process, research into modelling fluvial suspended sediment transport is still in a stage of comparative infancy. This is in part due to the difficulties in obtaining rainfall, flow and SSL data of sufficient quality, but also due to the underlying complexity of the many nonlinear processes that combine to control catchment sediment delivery. For example, the supply of sediment available for

transport is dependent upon the action of catchment hillslope processes, such as rill and gully development and mass movements (Selby, 1993). Further, the sediment transport capacity of a stream is strongly flow dependent, which in turn is an inherently nonlinear process. It follows, therefore, that a model must account for these nonlinearities to successfully characterise catchment sediment transport.

Current methods of prediction fall into two main categories. Firstly, deterministic models such as CREAMS (Knisel, 1980) and ANSWERS (Beasley *et al.*, 1980) that aim to capture the physical complexity of sediment supply and delivery processes. These deterministic models are often over-parameterised and require substantial field data for calibration. Secondly, simple nonlinear regression models between flow and SSL are often unable to adequately reproduce the dynamic nature of fluvial systems and are, therefore, really inappropriate for modelling suspended sediment (Walling, 1977). The short-comings of both techniques can be addressed, to some extent, by using transfer functions which are both dynamic and parsimonious in nature. Sharma *et al.*, (1979), Sharma and Dickinson, (1980), Lemke (1990; 1991) and Wang *et al.*, (1991), for example, successfully utilise linear transfer functions for modelling suspended sediment load where flow is used as the model input.

Using data obtained from 5 catchments situated within the Loess Plateau of China, Wang *et al.*, (1991) successfully adopted linear TF's to model suspended sediment load. Due to the particular physiographic characteristics of the Loess Plateau, sediment yield in this region is strongly transport dependent. The majority of rainfall in this region, derived from high magnitude short duration storms occurring in the summer months, is converted directly into runoff, with little lost to groundwater. As a

result, the relationship between runoff and sediment yield is very strong and a linear TF is sufficient to characterise the dynamic. Sharma *et al.*, (1979), Sharma and Dickinson (1980) and Lemke (1990; 1991) both identified a linear TF model between log transformed runoff and log transformed SSL time series. As a consequence of log transforming the data, they found that the nonlinear relationship existing between the two series had been partially linearised, to the extent that a linear TF was satisfactory to explain the behaviour within their study catchments.

The Wyresdale catchment is particularly flashy and the rainfall-flow and flow-SSL relationships are inherently nonlinear. Although a linear model is able to approximate the rainfall-SSL relationship during wet conditions, it is unable to fit the data well during periods of low flow. Consequently, for the purposes of reconstructing the catchments SSL series, a hindcasting model with a nonlinear structure is required.

5.3.1 Wyresdale Rainfall-Sediment model

Following the DBM approach outlined previously, the relationship between the Wyresdale 1994-1996 rainfall and SSL was investigated using a time varying parameter approach. Non-parametric estimation, utilising the FIS algorithm, indicates that the numerator parameter $\hat{b}_0(k \mid N)$ exhibits significant temporal variation. On inspection, the TVP showed a satisfactory dependence on flow and rainfall, but state dependent parameter modelling (SDPM) was unable to identify a sufficiently clear and reliable state dependent relationship. Nevertheless, a satisfactory relationship that is able to approximate the key nonlinearity of the rainfall-sediment system, generating a *transformed rainfall* series (u_{se}), is presented in equation (5.4) and forms the first

component of the model. The general rainfall-sediment model was identified as having the following structure,

$$u_{se}(k) = u(k)^{\alpha}\, y(k)^{\beta} + \zeta(k) \tag{5.4}$$

$$s_i(k) = H\, u_{se}(k-\delta) + \varepsilon(k) \tag{5.5}$$

where, $s_i(k)$ is the mean daily instantaneous SSL at the reservoir inflow (kgs^{-1}); $u(k)$ is daily rainfall (mm); $u_{se}(k-\delta)$ is the transformed rainfall with pure time delay δ; $y(k)$ mean daily instantaneous flow at the reservoir inflow (m^3sec^{-1}); α, β and H are constant coefficients; and $\zeta(k)$ and $\varepsilon(k)$ are general noise terms included to account for stochastic disturbances and unmeasured inputs to the system.

The Wyresdale rainfall-sediment model is similar in principle to the rainfall-flow model of Young & Beven (1994), but with the rainfall term, $u(k)$ raised to a power, which is included to enhance intense rainfall events as a proxy indicator of greater erosive capability. The second part of equation (5.4) ($y(k)^{\beta}$) acts as a measure of the catchment's antecedent soil moisture conditions, characterising the nonlinear rainfall-flow effect. Overall, the two combined nonlinear components of the transformed rainfall equation (5.4) can be interpreted as a low pass filter, smoothing the fluctuating rainfall and producing a measure of the rainfall that is said to contribute to the movement of sediment through the catchment. Essentially, this term aims to characterise the complete nonlinear behaviour of the system in the same manner as the *effective rainfall* term of rainfall-flow models.

For the Wyresdale data series, the effect of daily averaging, combined with the known rapid response of the catchment, reduces the dynamic content of the data. Therefore, for equation (5.5), only a simple linear model with a gain parameter H is required, between the transformed rainfall (u_{se}) and the suspended sediment load (s_i) at the reservoir inflow. In circumstances where the available data is more frequently sampled and/or the dynamic content of the SSL series is greater, the simple gain parameter H would normally be replaced by a TF. The model parameters in equations (5.4) and (5.5) are estimated utilising an iterative nonlinear least squares optimisation routine, incorporating the SRIV estimation algorithm. The structure of the optimisation algorithm is shown in Figure 5.5. The best rainfall-sediment model identified in this manner for the 1994-1996 Wyresdale series is first order, which explains 91% of the data ($YIC = -7.687, R_T^2 = 0.912, AIC = -10.718$). The parameters estimated are $a_1 = -0.171(0.024)$, $b_0 = 0.0021(5.495e^{-5})$, $\alpha = 0.40(0.01)$ and $\beta = 0.75(0.01)$, where standard errors are given in parentheses.

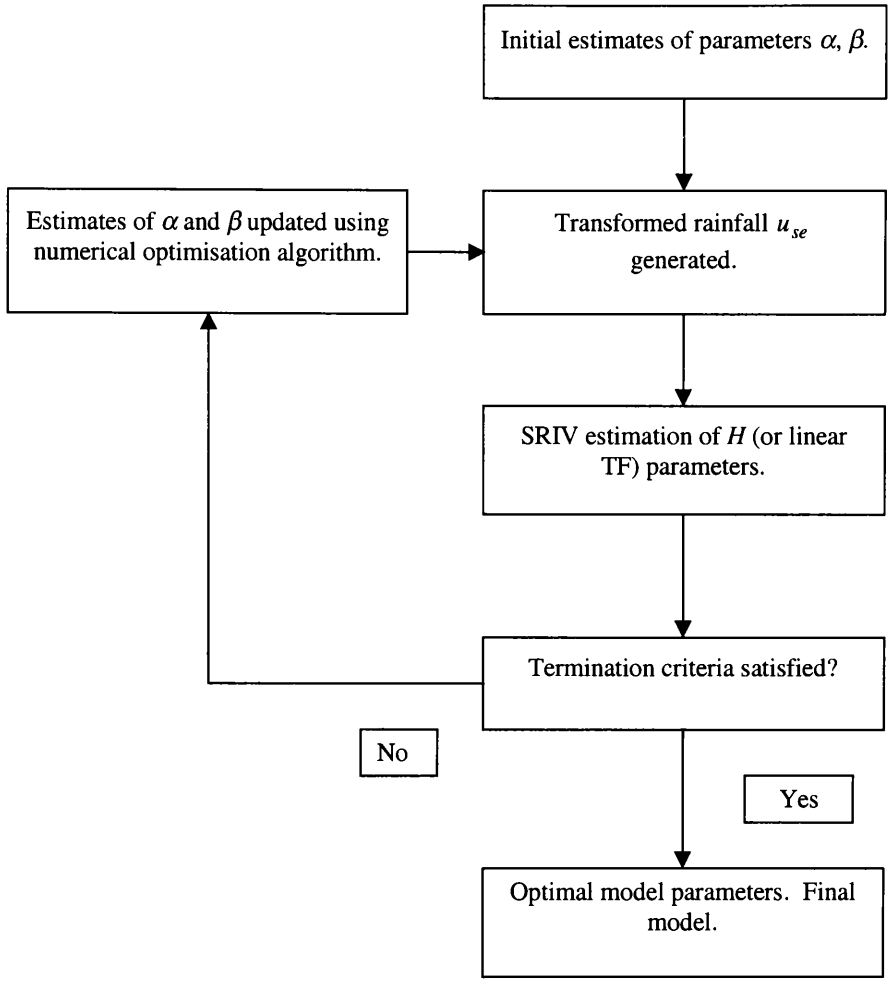

Figure 5.5 Schematic of the parameter optimisation procedure.

The rainfall-sediment model (equations 5.4-5.5) accurately fits all of the major storm peaks of the reservoir inflow SSL series and also manages to fit the majority of the lower SSL conditions, as shown in Figure 5.6. In a similar manner to the rainfall-flow model identified in Sections 5.2.2, storm events following prolonged periods of warm dry weather or associated with convective precipitation, are not modelled as well as all other conditions. Overall, the respectable performance of the model, indicates that its structure is suitable for modelling the rainfall-sediment system.

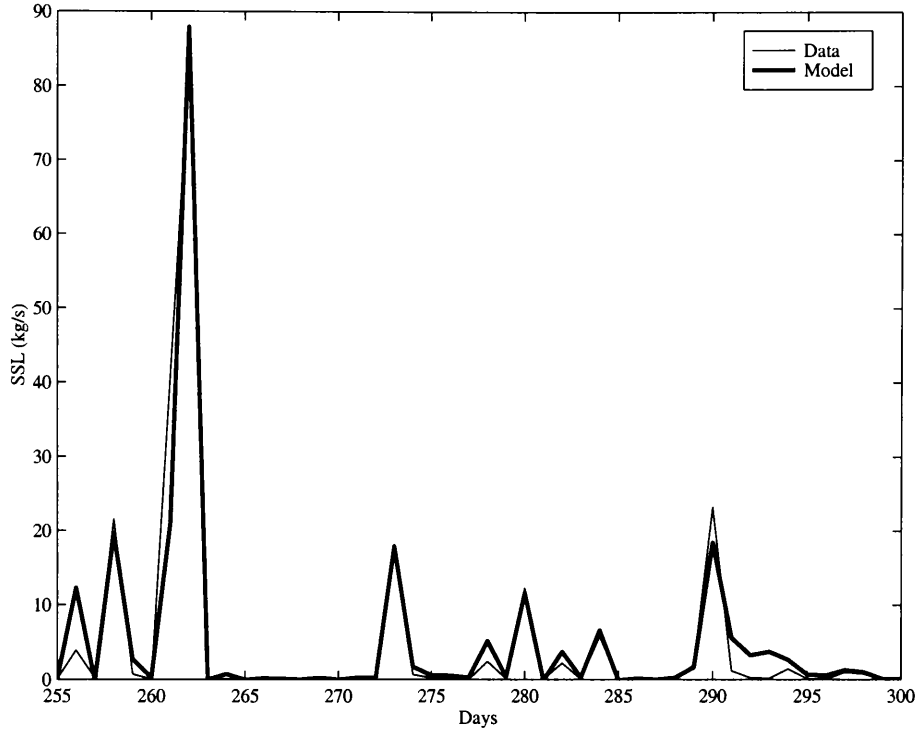

Figure 5.6 Nonlinear TF sediment model. Data (fine line) and model (bold line).

5.3.2 Wyresdale rainfall-sediment model for hindcasting

Whilst the rainfall-sediment model presented in equations (5.4) and (5.5) is suited for general modelling and forecasting applications, its structure is inappropriate for generating a synthetic sediment series by hindcasting. As only historical rainfall data are available for Wyresdale, the *transformed rainfall* equation (5.4) is modified to the following structure,

$$u_{se}(k) = u(k)^{\alpha}\, y(k)^{\beta} + \zeta(k); \qquad y(k) = \frac{1}{\hat{A}(z^{-1})} u(k) \tag{5.6}$$

where, $u(k)$ is daily rainfall (mm); $u_{se}(k-\delta)$ is the transformed rainfall with pure time delay δ; α and β are constant coefficients; and $\zeta(k)$ is a general noise term. The daily mean flow term $y(k)$ in equation (5.4) has been replaced by a linear TF

where the polynomial $A(z^{-1})$ is defined as $A(z^{-1}) = 1 + a_1 z^{-1} + ... + a_n z^{-n}$. The second nonlinear term accounts for the catchment antecedent soil moisture conditions in a similar manner to the antecedent precipitation index (API) (Shaw, 1988). The second component of the rainfall-sediment model remains unchanged, as presented in equation (5.5). Through an iterative procedure, the polynomial $A(z^{-1})$ was identified as first order and the parameters in equations (5.5) and (5.6) were estimated using the optimisation routine as before. The best estimated model parameters are as follows, $H = 0.0013$, $a_1 = -0.971(0.024)$, $\beta = 1.78(0.082)$ and $\alpha = 1.312(0.115)$ where the parameters standard errors are given in parentheses. The rainfall-sediment model fit now explains 84% of the data ($R_T^2 = 0.84$). This good fit, as may be seen in Figure 5.7, indicates that the model can be used with some confidence, as a simulation model for hindcasting sediment loads into the reservoir.

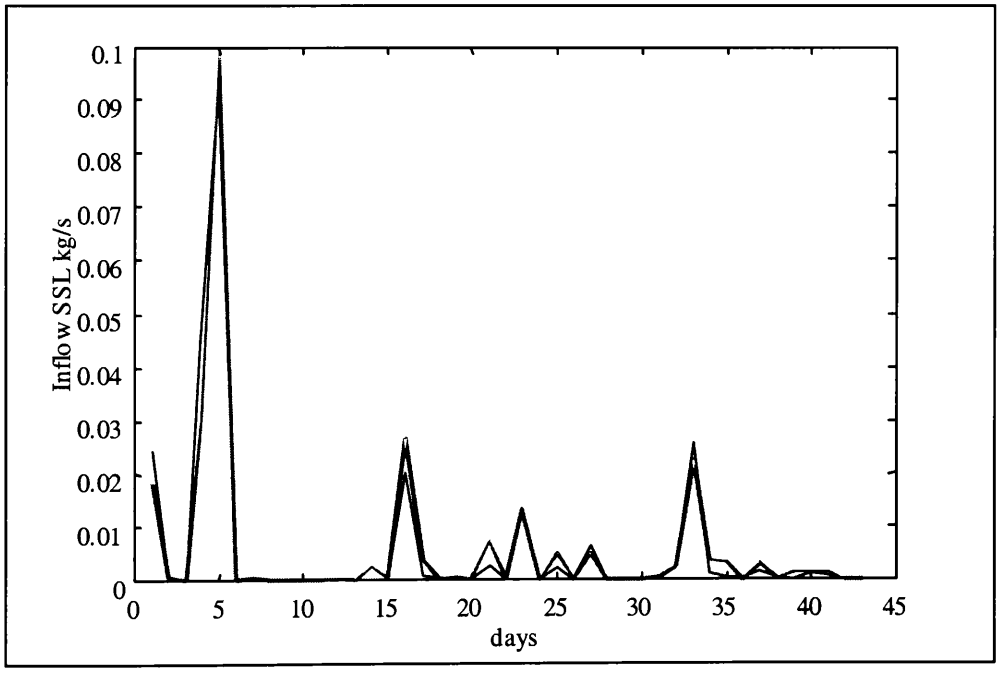

Figure 5.7 Model fit of the hindcasting model: data (fine line) model (bold line)

5.3.3 Reservoir sediment transmission model

Having identified a model to predict the input of SSL into the reservoir, the second component in this DBM modelling scheme is a very simple nonlinear TF model, estimated between the mean daily SSL flux (kgs^{-1}) at the reservoir inflow (s_i) and the mean daily SSL (kgs^{-1}) at the reservoir outflow (s_o). Utilising the SRIV algorithm, the best model for the transmission of the sediment through the reservoir was identified as having the following structure with standard errors given in parentheses,

$$s_0(k) = \frac{1.952(0.025)}{1-0.149(0.011)} s_i(k)^{1.178} + \xi(k) \tag{5.7}$$

where $s_0(k)$ is mean daily sediment flux at the outflow; $s_i(k)$ is mean daily sediment flux at the inflow; and $\xi(k)$ is an additional general noise term. This model explains 92% of the data with a coefficient of determination, $R_T^2 = 0.918$.

5.4 GENERATING HISTORICAL SERIES

The combined rainfall-SSL model (equations 5.5-5.7) calibrated over 2 years of data, may now be used to extend the sediment records at the reservoir inflow and outflow by running the available rainfall sequence from 1911-1996 through the complete model. At the commencement of the simulation, the model is initiated without any prior 'knowledge' of the catchment soil moisture conditions. Consequently, a '*run in period*' is required to ensure that the transformed linear TF, which describes the antecedent condition component of the model, is stable (see equation 5.6). With each subsequent k*th* time step of the simulation, a picture of the soil moisture conditions is

built up which gradually becomes stable. The time period required to ensure such stability is dependent upon the time constant of the linear TF and the β coefficient, which, in this instance, is 150 days.

Simulating the complete model in this manner, produces the 86 year, daily flow and sediment sequences presented in Figure 5.8. Further validation of the model can be achieved by comparing the last period of the synthetic SSL outflow series for 1994-1996 with the observed field data for the same period. The overall fit remains very high with a coefficient of determination R_T^2 over this period of 0.91. The effects of the different antecedent conditions are clearly evident in Figure 5.8: for example, the propagation of two equal magnitude rainfall events through the model can result in highly contrasting SSL outputs. When comparing the complete 86 year rainfall and simulated SSL output, it is clear that the model has preserved the catchments SSL dynamics, limiting the majority of SSL transport to a small number of low frequency high magnitude events. Additionally, the significance of low recurrence, high magnitude events can be evaluated by this approach: the SSL load for each event can be estimated from the series and evaluated in relation to the estimated annual inflow load. For example, the 3 storms 27-28th October and 15th, 20th November 1980 are estimated, from the model, to collectively contribute 31% of the simulated sediment input for that year.

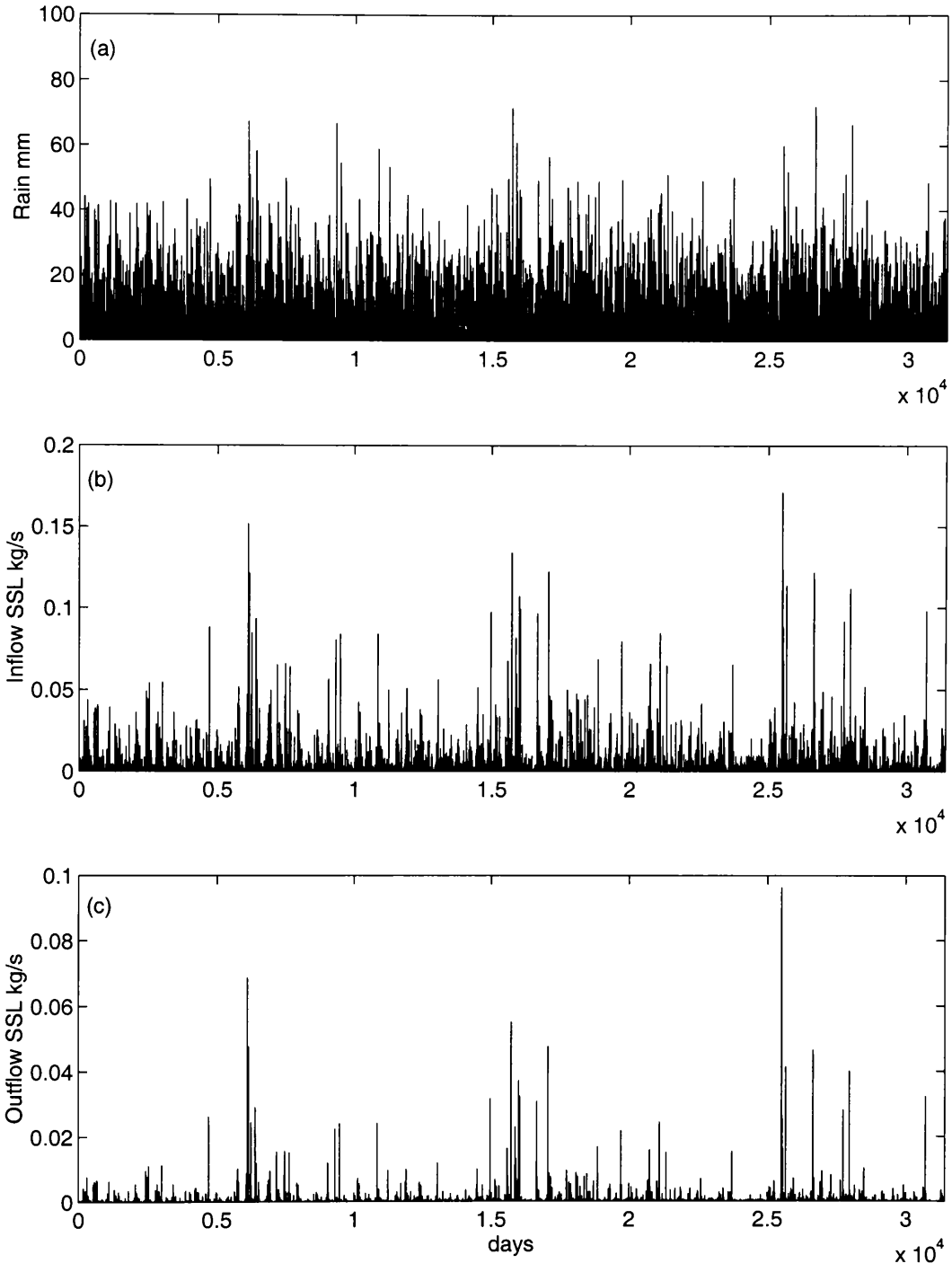

Figure 5.8 86 year (a) rainfall series, (b) simulated SSL into the reservoir and (c) simulated SSL out of the reservoir.

From the simulated series the total deposited suspended sediment load in 86 years was estimated to be 3096 tonnes which, when after an approximate bedload component is added (c.f Goodwill, 1998), gives an average deposition of 57 tonnes y^{-1}. Although

this is higher than the 50.1 tonnes y^{-1} estimated from the bathymetric survey, the two figures are encouragingly close, considering the uncertainties involved in generating both the modelled series and conducting the bathymetric survey.

5.4.1 Monte Carlo uncertainty analysis

Since the overall model, represented by equations (5.5-5.7), is calibrated from only two years of experimental data, a measure of uncertainty should be assimilated. Fortunately, the stochastic nature of the DBM methodology and SRIV estimation procedure enables the uncertainty associated with the parameters to be incorporated within a Monte Carlo (MC) framework (e.g. Whitehead & Young, 1979).

The cumulative sediment load graph for the 1994-1996 Wyresdale data is presented in Figure 5.9. It is shown that 85% of the outflow SSL is transported in only 2% of the time, which is indicative of a flashy catchment. As a means of assessing the model uncertainty, a series of 3000 MC simulations were run using the historical 86 year rainfall series as the model input. The model cumulative sediment load curve for the 1994-1996 two year period, extracted from the synthetic series, with standard error bounds calculated from the ensemble of MC simulations, are also plotted on Figure 5.9. The shape of the synthetic curve is similar to that of the measured series, which resides within the model standard errors. Moreover, the curves shown in Figure 5.9 provide further evidence that the model is capturing the overall dynamic of the catchment sediment response reasonably well.

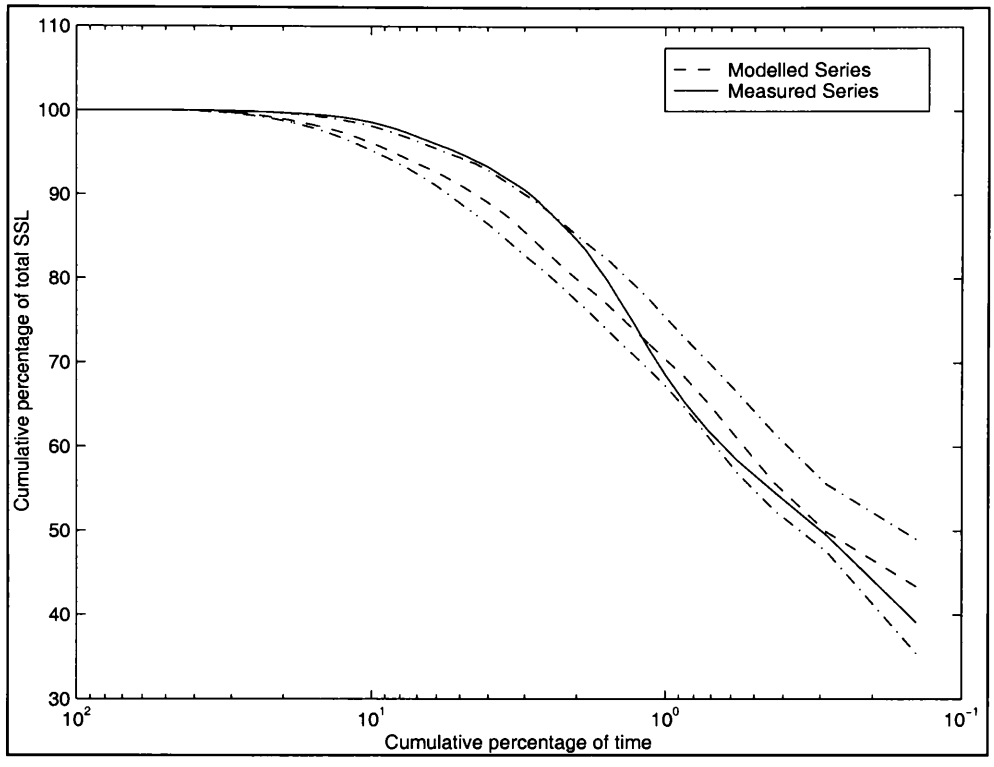

Figure 5.9 Cumulative sediment load curve for the 2 year (1994-1996) measured outflow series (full line) and cumulative sediment load curve for the same two year period from the historical 86 year outflow sequence (dashed) with standard error bands calculated from the 3000 MC simulations (dot-dashed).

5.5 VALIDATION OF THE MODEL

Through the DBM modelling process outlined in the previous section, a synthetic 86 year inflow and outflow SSL sequence has been generated. The similarity between the 2 year measured field data and the corresponding section of the synthetic sequence, suggests that the 86 year inflow and outflow series can be used to provide a reasonable estimate of sedimentary deposition in the reservoir. In an attempt to validate this assertion, sediment cores were extracted from Wyresdale reservoir to compare directly with the reconstructed synthetic sediment accretion sequence.

5.5.1 Synthetic sediment accretion reconstruction

An estimate of the daily sediment deposition rate in the reservoir between 1911-1996 may be obtained by subtracting the outflow sequence from the inflow sequence. The series of daily sediment load (kgs^{-1}) settling in the lake can then be converted into a synthetic sediment accretion sequence by distributing this load throughout the main (*c.* 27 000 m^2) depositional basin on the reservoir bed. A proxy indicator of spatially averaged sediment accretion can then be obtained by dividing this sediment sequence by the average bulk density of the Wyresdale sediments (0.242g cm^3). In this manner, the sediment load over the 86 years is converted into an estimated sediment depth of 48 cm, which is equivalent to an average annual accumulation rate of ~0.55 cm y^{-1}. The resulting synthetic sediment accretion sequence, shown in Figure 5.10, shows a number of high yearly deposition rates, which should correspond to discrete deposition horizons in the reservoir if the model is representing the sedimentary depositional characteristics reasonably.

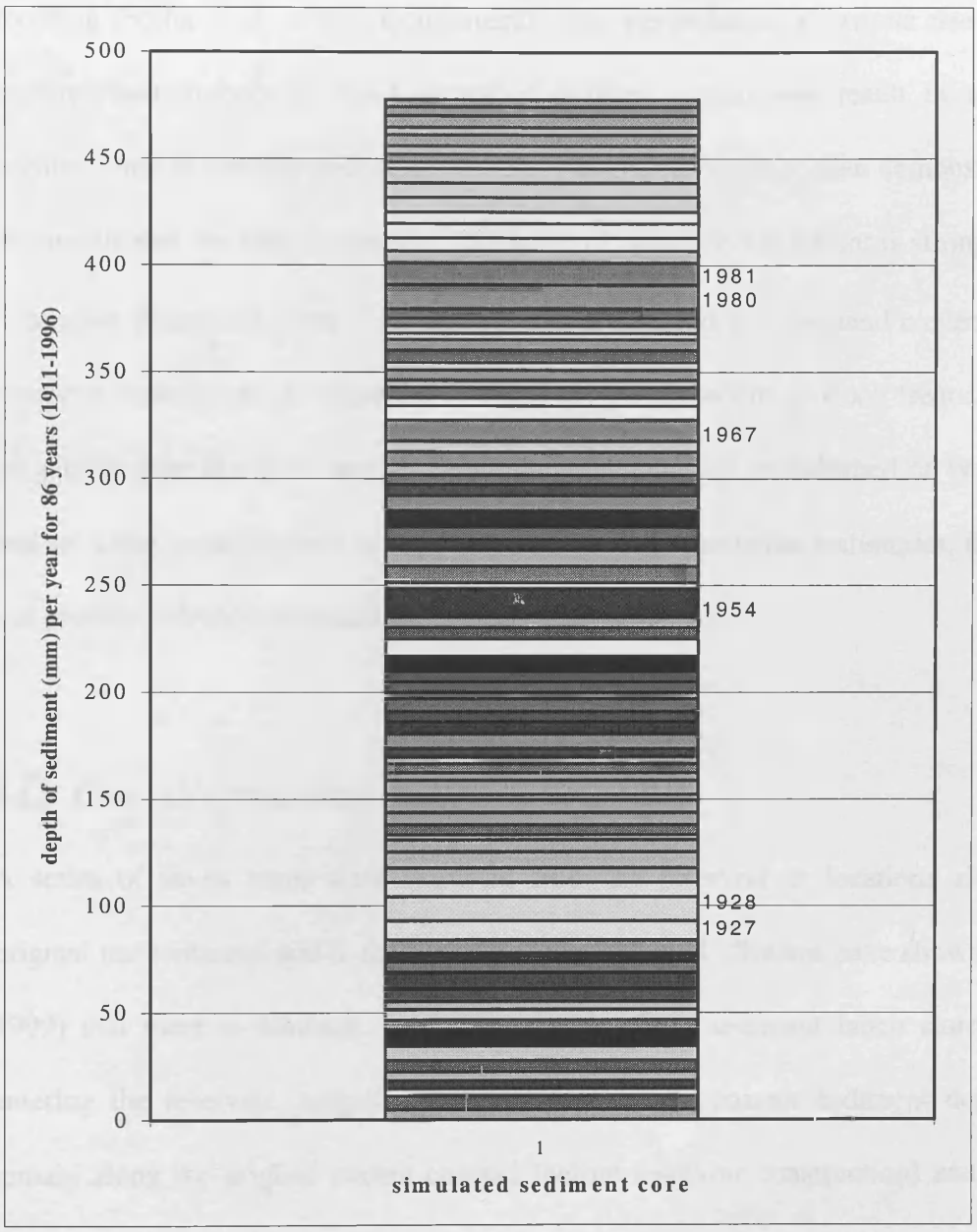

Figure 5.10 Simulated sediment core profile (48cm deep in total) Depth of sediment deposited each year in millimetres comprises a separate band in the profile(each year is identified by a different colour).

5.5.2 Lake sediment cores

Reservoir sedimentation rates typically exhibit spatial and temporal variations, with additional complexities associated with sediment fining, re-suspension and sediment

focusing (Foster *et al.*, 1986). Consequently, the interpretation of particle size profiles requires caution because flood events of different magnitudes result in sediment accumulation at variable rates (Curr, 1985). Nevertheless, it has been demonstrated at Wyresdale that the sand component ($> 63\mu m$) of suspended sediment is strongly flow dependent (Goodwill, 1998). As such, it can be inferred that the sand content of the reservoir deposits should represent an important proxy record of flood frequency and magnitude over the study period. The lake sediments are undisturbed at Wyresdale and so, using a combination of sediment analysis and correlation techniques, the cores can provide valuable information for model validation.

5.5.3 Core extraction and analyses

A series of seven cores were extracted from the reservoir at locations along the original main channel and in the vicinity of the dam wall. Studies have shown (Price, 1999) that there is minimal dispersive mixing of the sediment laden storm water entering the reservoir, suggesting the main zones for coarser sediment deposition remain along the original stream channel (before reservoir construction) and behind the reservoir dam. Cores were extracted in the reservoir from a platform of inflatable boats, as shown in Plate 5.2a, following a technique developed by Mackereth (1969) which enables an *in situ,* undisturbed record of sediment to be collected. The apparatus comprises of an open ended plastic cylindrical core that is forced into the basal sediment by compressed air. Of the seven sediment cores extracted, cores 1 and 5 were considered to show significant signs of disturbance and were excluded from further analysis.

Plate 5.2 (a) Platform of boats and coring equipment at Wyresdale reservoir (left) and (b) the extracted sediment cores (right).

When selecting a single core (Plate 5.2b) for detailed particle size and radioceasium analysis, it is important to ensure that it is representative of the sediment depositional characteristics of the reservoir. Magnetic susceptibility assesses the degree to which a material can be magnetised (Foster *et al.*, 1986), expressed per unit volume of the sample (κ), which is directly related to the concentration of ferromagnetic grains within the materials composition. In the laboratory, the *in situ* volumetric magnetic susceptibility of each sediment core was determined using a *Barlington MS2C* magnetic susceptibility loop and the *Multius* software. The magnetic susceptibility profiles for four of the sediment cores are presented in Figure 5.11. The similarity between the shape of the profiles indicate that representative cores have been extruded from the reservoir. Core 2 was selected for further analysis because on close examination, the colour and size of the sediment at the base of the core suggested that the pre-inundation surface had been captured, indicating that a complete record of depositional sediment had been acquired.

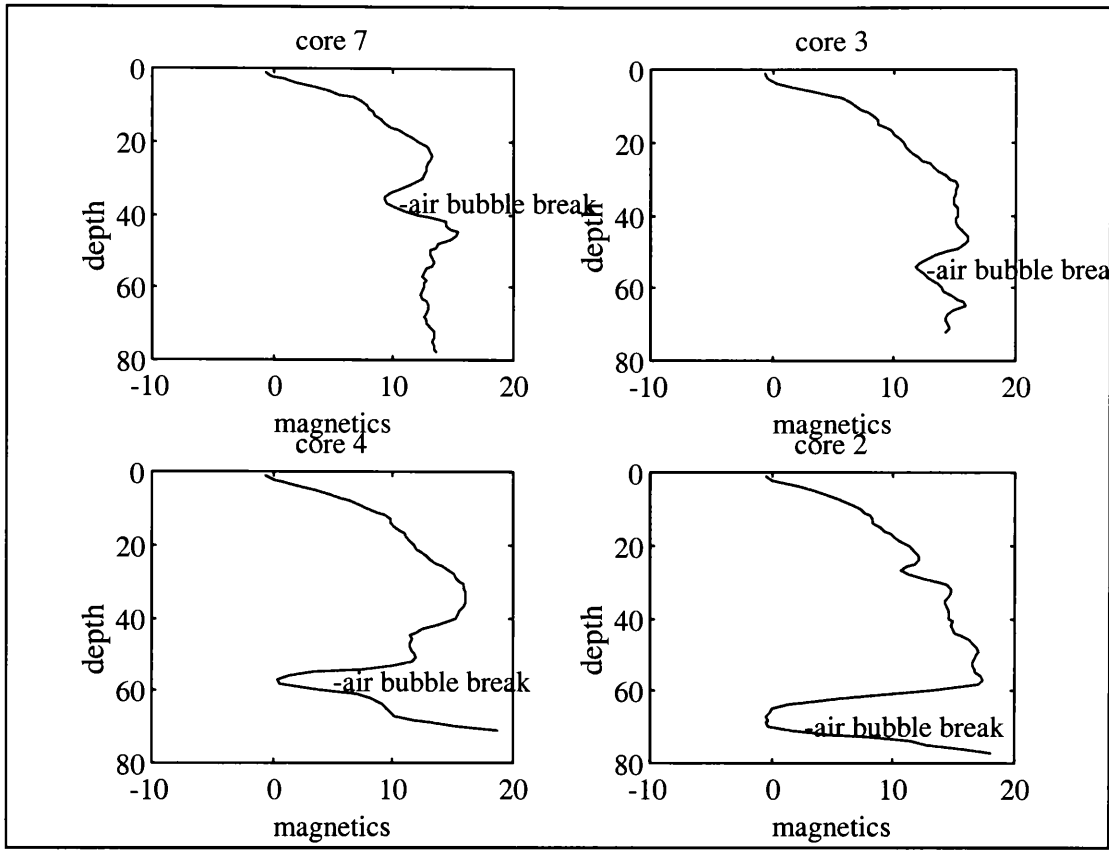

Figure 5.11 Magnetic susceptibility of four sediment cores extracted from Wyresdale reservoir.

The particle size distribution of 0.6 cm increments of the core were measured using the Coulter LS-230 laser particle sizer. The equipment consists of three components: a fluid unit used to introduce the sediment into the equipment for analysis; the laser particle sizer; and a PC which runs the software controlling the operation. The particle size distribution of the sediment sample is determined by measuring the laser diffraction patterns of sediment particles over a size range of 0.04-2,000 μm. Prior to analysis, $10cm^3$ of 0.4% Calgon was added to the sediment to induce dispersal without peroxide digestion of organic material (Duck,1996) and each sample was placed on a shaker overnight.

The profiles presented in Figure 5.12 show the relative proportions of sand, silt and clay throughout the length of the core. Figure 5.13 presents the vertical profiles of the core mean wet and dry bulk density. The clay fraction remains approximately constant at a low percentage throughout the length of the core, whilst the sand fraction increases and silt fraction decreases towards the base. Sharp increases in the percentage sand fraction and the dry bulk density can be noted at a number of discrete intervals along the core length. As the sand component of the SSL at Wyresdale has been found to be flow dependent (Goodwill, 1998), these isolated, denser horizons are indicative of high magnitude storm events. At a depth of below 54 cm, there is a marked increase in the sand content. Microscopic examination of this coarse sandy material showed it to be well-sorted, consistent with sediment deposited in a fluvial environment, i.e. a pre-inundation environment (Dr. J. Rowan *pers. com.*).

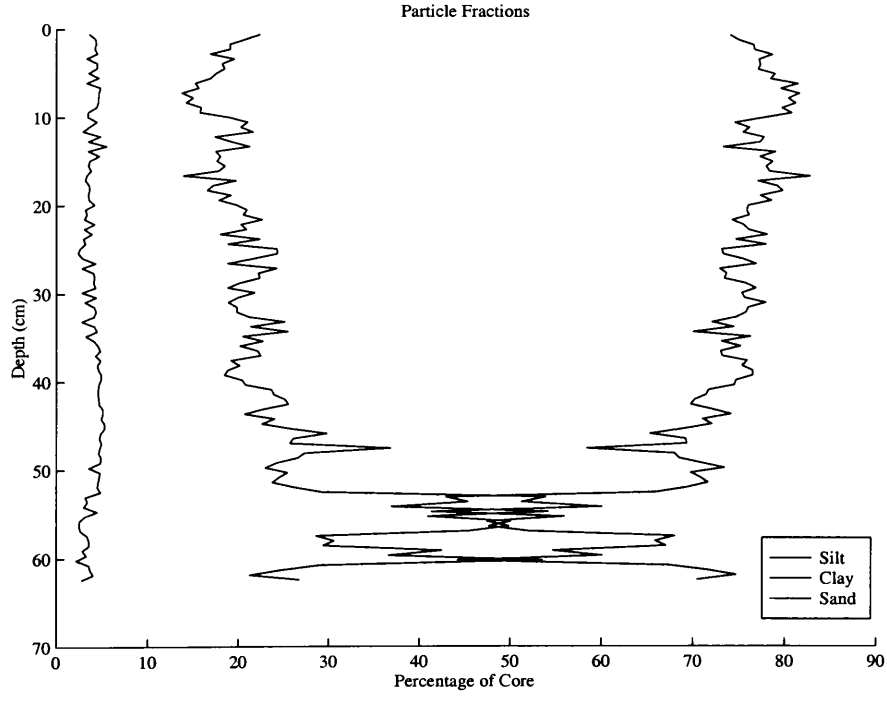

Figure 5.12 Percentage content of clay (left profile), sand (middle profile) and silt (right profile) in Core 2.

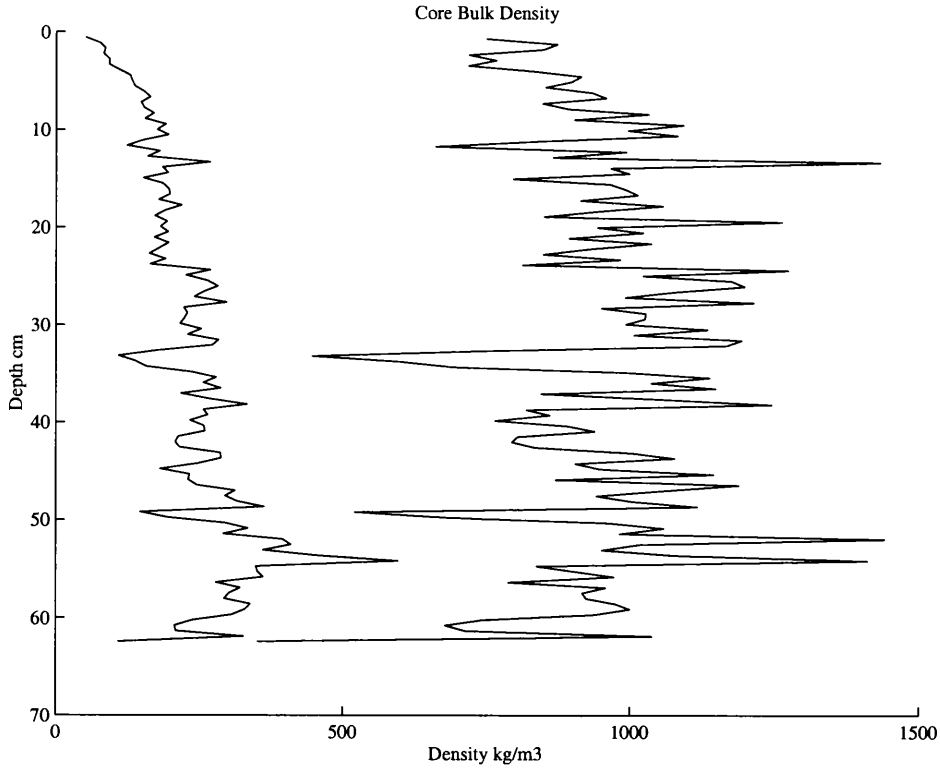

Figure 5.13 Wet bulk density (right profile) and dry bulk density (left profile) of Core 2.

5.5.4 Core validation

In order to assist a comparison of both profiles, the core was independently analysed for ^{137}Cs using gamma spectrometry to generate an absolute dating chronology. A sharp peak in ^{137}Cs level is observed at a depth of 8-10 cm, which is considered to be attributable to the 1986 Chernobyl disaster. A broader peak is located at a depth of 18-20cm which is caused by above ground atomic weapon tests which occurred between 1959-1963 (c.f. Rowan *et al.*, 1995). No further ^{137}Cs was detected below this depth. The resultant chronology from this analysis is shown in Figure 5.14, at the corresponding depth of sediment.

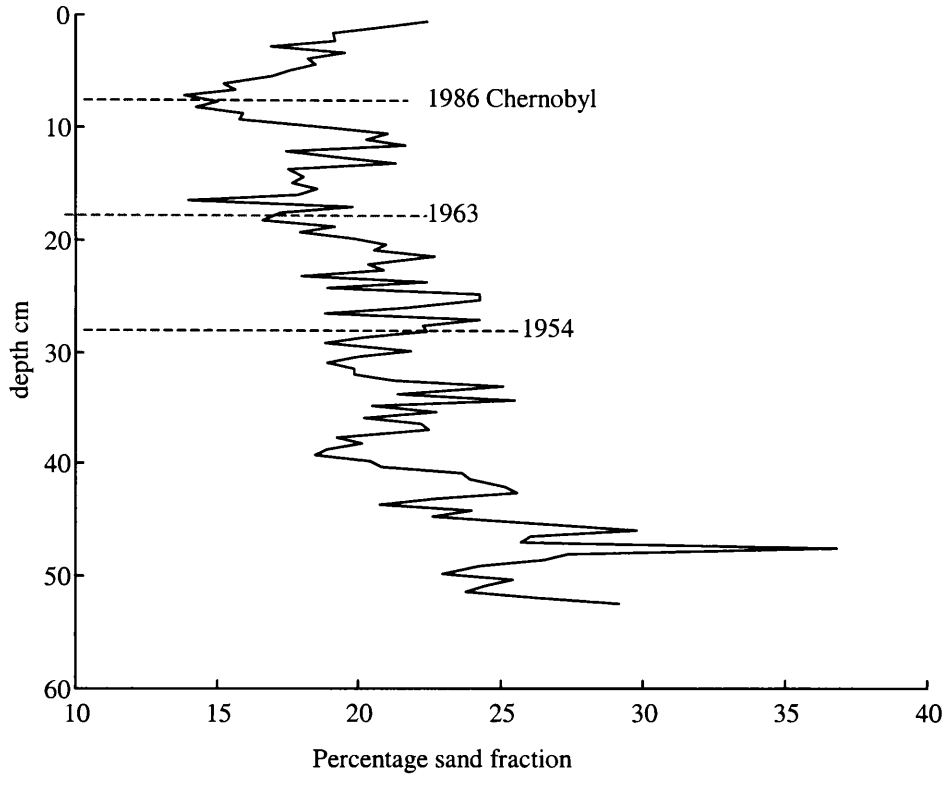

Figure 5.14 Percentage sand content of Core 2 with ^{137}Cs dating chronology.

It is not possible to quantitatively compare the synthetic sediment accretion sequence, which is expressed in daily deposition load, with the sand fraction profile of the extracted sediment core. However, it has been demonstrated at Wyresdale that the sand component of the SSL is strongly flow dependent. As a result, a qualitative comparison of the sediment core and synthetic profile can be made by plotting the two sequences adjacently and matching peaks and troughs assisted by the ^{137}Cs dates. As shown in Figure 5.15 the two profiles show good visual agreement.

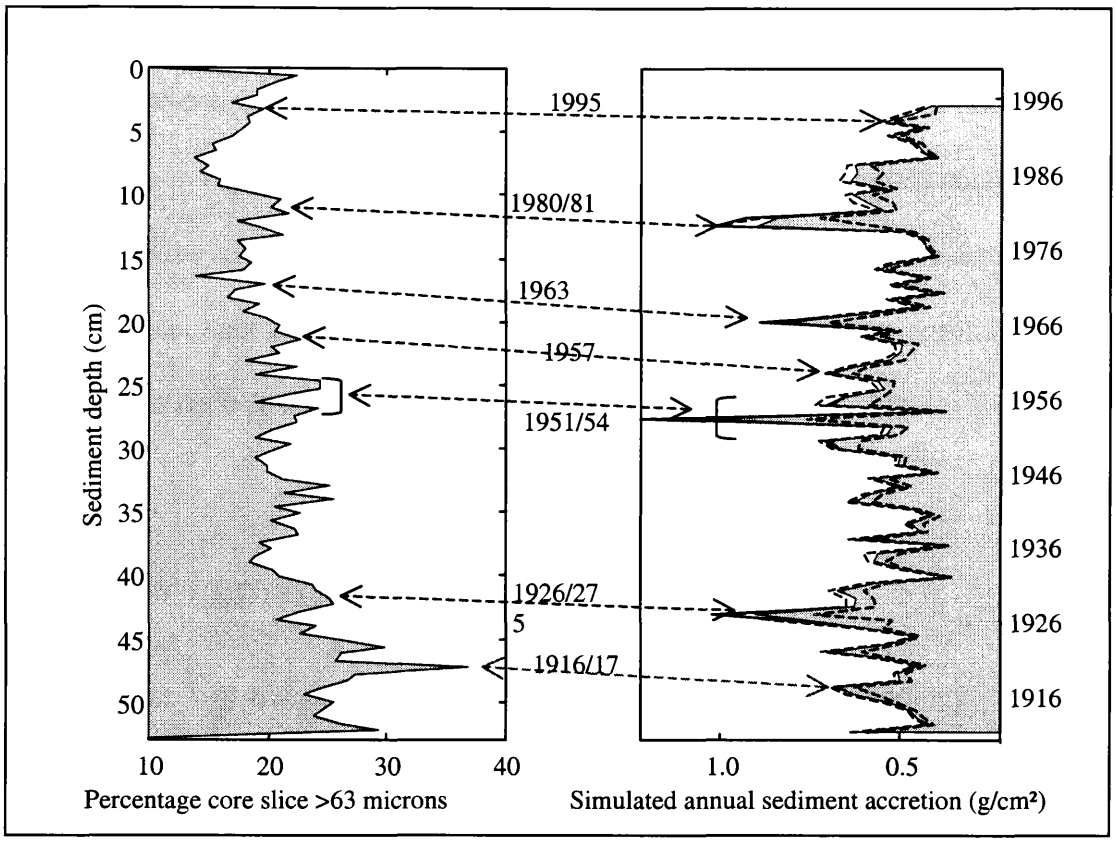

Figure 5.15 Simulated and observed accretion sequence

Of particular interest are the high rates of sediment deposition simulated by the model during the winters of 1916/1917, 1927, 1954, 1963 and 1980/81. There appears to be corresponding evidence of significant deposition horizons in the lake at these time points, reflected by the increased percentage sand fractions in the core. The significance of the larger storms predicted from the model are additionally confirmed from independent evidence corroborated from local newspaper records. For example, on 28th October 1980 the model estimated that 6.5 tonnes of sediment was deposited in the reservoir. On 31st October 1980 the Garstang Guardian headlined with 'Water up to 4ft in places'; 'M6 closed for several hours', including making specific references to Scorton village. The model estimated that 6.45 tonnes of sediment were deposited on 17th September 1957 and subsequently on 20th September 1957 the

Lancaster Guardian reported high flood water levels with the headline: 'Wye overflow'; 'Scorton road to A6 closed'. The analogous trends identified between the two profiles in Figure 5.15 indicate that the DBM models utilised in this analysis have to some degree captured the dominant sediment dynamics of the catchment-reservoir system and an encouraging reconstruction of the historical sediment accretion in the reservoir has been possible.

5.6 CONCLUDING REMARKS

This chapter has discussed the application of the DBM methodology to model historical sedimentation sequences at Wyresdale Park catchment. A novel nonlinear sediment model has been used to simulate suspended sediment loads from rainfall records alone, providing information on sediment transport dynamics over much longer periods than are currently available in most monitoring programmes.

Initially, an investigation of the rainfall-flow dynamics at the Wyresdale catchment was made to help in the development of a rainfall-sediment model. Rainfall-flow modelling at the seasonal scale showed evidence of some time varying behaviour within the model parameters such that a single model with seasonally adjusting parameters could potentially be estimated. However, as these relationships were not clearly defined, a model of this complexity was deemed to add unnecessary uncertainty. As a result, the rainfall-flow and rainfall-sediment processes at Wyresdale, were modelled based upon the complete two years of time series data. The identified rainfall-sediment model is similar in principle to the rainfall-flow model. It is comprised of a nonlinear and linear element: the nonlinear component provides a

measure of the rainfall and flow that contributes to the transport of the sediment through the catchment system; and a second linear TF which relates this transformed rainfall to suspended sediment load.

This preliminary study has demonstrated the potential of the DBM methodology for producing historical time series. This is increasingly important where SSL data is unavailable in sufficient quantities and where it is expensive to collect. These historical data can be used to explore magnitude-frequency relationships and their implications for event horizon preservation within sediment profiles. The TF model used to generate the synthetic sediment sequence assumes that the dynamics of the contemporary catchment are stationary and can be extended over historical time-scales. However, the sand fraction of the sediment core and the average dry bulk density increase below a depth of 20 cm as shown in Figures 5.13 and 5.15, which may suggest the system is actually time variant. The preliminary results of this study indicate that the modelling scheme is able to reproduce the calibration data in the sense that the model has been able to predict the location of a number of discrete flood horizons. However, the increasing trend in the sand fraction and bulk density of the sediments can not be explained by the model. The deviation of the observed data from the stationary model predictions, particularly in the lower portion of the profile, suggests that the behavioural response of the catchment system may have changed, possibly due to land use or land management modifications. Although the results obtained in this study are not conclusive, it represents the first such exercise of its type, and in this regard, the results are very encouraging.

The original premise of the experimental programme at Wyresdale was not intended to produce data for rainfall-flow or rainfall-sediment modelling. Nevertheless, the DBM approach has demonstrated for the first time that even with limited data it is possible to model the nonlinear rainfall-sediment relationship. It is hoped that this DBM approach could be developed further for other better suited catchments for use as an additional design tool for water resources management. Ideally, to explore the full potential of this DBM approach, it would require a catchment with extended time series which exhibit greater dynamics, i.e. sampled more frequently e.g. every 15 minutes, or choosing a larger catchment where daily sampling would be adequate.

The work summarised in this chapter highlights how the DBM modelling methodology can be applied to the identification and estimation of nonlinear hydrological models. Future work aims to look more closely at the nonlinear transformed rainfall term which lumps the sediment production processes through FIS/SDPM analysis. The study also provides a good platform for other research initiatives, such as evaluating the effects of model uncertainty and the implementation of Kalman filter-based forecasting models for both discharge and sediment transmission.

CHAPTER 6

OCEANIC ECOSYSTEM MODELLING

Plankton play an important role in oceanic-atmospheric dynamics at different spatial and temporal scales through many biophysical processes (Fasham, 1993). As a consequence of increases in anthroprogenic gases emitted into the atmosphere, particularly CO_2, there is increasing concern that the earth may experience global climate change within the next century. The oceanic carbon cycle is a key component of global carbon cycling and it is becoming ever more important to quantify both the uptake and effects of anthroprogenic CO_2 in the ocean. Marine biota play a significant role in carbon cycling, but it is assumed that these biota will not be affected directly through uptake of anthroprogenic CO_2 (Sarmiento *et al.*, 1992). However, global climate induced changes to the physical processes affecting the marine biota may have a significant impact on the ecosystem. Understanding the oceanic ecosystem, particularly how the key elements of phytoplankton, zooplankton and nutrients interact, is vital, especially as phytoplankton and zooplankton play fundamental roles in the marine food web supporting a diverse range of terrestrial and marine life. It is essential that mathematical models are used as an effective tool to help fully understand the plankton population dynamics and the interaction of the physical

processes which drive the dynamics of this system (i.e. mixed layer depth, water temperature, total radiance), particularly as these dynamics may alter as a result of global climate change. The ability to accurately quantify how the ecosystem may react to climate modifications is essential, as changes in total biomass or its spatial and temporal distribution in the ocean could have disastrous effects on human life (SECOR (1990).

6.1 CHAPTER AIMS AND OBJECTIVES

The oceanic ecosystem, as with most environmental systems, is governed by extremely complex dynamic processes which are inherently nonlinear. Also, even after extensive scientific investigations over the past 20 years, significant uncertainty still remains about its general nature and structure.

This chapter presents an evaluation of a nonlinear, deterministic ecosystem model in two main components. Firstly, parameter optimisation and the effects of parametric uncertainty propagating through the model are investigated. Then genetic algorithm and nonlinear least squares parameter optimisation techniques are applied to the ecosystem model. In these studies, the results of exploring the parameter hyper-space provide initial estimates of parameter sensitivity that can be utilised in a more thorough evaluation of model uncertainty following the data-based mechanistic DBM modelling philosophy. The deterministic simulation equations of the system are considered in stochastic form by assuming that the associated parameters of the model are uncertain. Monte Carlo Simulation (MCS) (see e.g. Whitehead and Young, 1979; Young *et al.*, 1996) and the associated Generalised Sensitivity Analysis (GSA) (Spear

and Hornberger, 1980) are utilised to explore this uncertainty, with the aim of identifying the parameters in the stochastic model that are most important in representing the dominant dynamic behaviour.

The second component of this chapter considers the Kalman filter (Kalman, 1960) and associated Maximum Likelihood (ML) method (see e.g. Harvey, 1989) as a further optimisation tool that can also be considered as a prelude to the development of a basic *data assimilation* method. Finally, by utilising the stochastic simulation model as a suitable surrogate for the actual system, a preliminary study based on combined statistical linearisation and model order reduction is undertaken, to identify whether a reduced, low order model can describe accurately the dynamic behaviour of the higher order system.

6.1.1 Data assimilation

The oceanographic system is extremely complex and it can be extremely expensive, difficult and sometimes impossible to obtain direct accurate measurements of oceanographic variables on a substantial and sustained basis for scientific and management purposes. Data assimilation refers to the melding of observational data with a model that characterises the dynamic principles governing the system at *run time*. In this regard, data assimilation is a powerful tool that can be used to help alleviate such problems of data acquisition by providing efficient, accurate and realistic estimations of important variables that often can not be feasibly measured, based upon the available observational data and the dynamic model. Data assimilation takes its roots in meteorology and engineering, but has more recently entered the field

of oceanography where it is expected to provide rapid advancements in basic ocean science and marine technology. The data assimilation framework consists of three components: observations, a dynamic model and an assimilation or melding scheme. The available observations and the dynamic model are assimilated via the melding scheme to provide estimates of the states or 'fields' of interest.

Data assimilation has wide application but has proved particularly useful in the following areas (Robinson *et al.*, 1999). Firstly, data assimilation can be used to minimise the predictability error of a model by continually updating the model's parameters/structure as new data becomes available so that reliable and accurate forecasts are maintained. Secondly, where the dynamic model is deficient, data assimilation can be used as a tool to continually adjust the model using the available data, such that it provides reasonable quantitative estimates of the state variables of interest. Thirdly, it can be used for state reconstruction: data can be assimilated into the model such to provide estimates of states that can be directly measured. Fourthly, to gain a better understanding of the dynamic processes of the system under consideration. Finally, the dynamic model and simulated high resolution data can be used to determine where key experimental networks should to be located such to increase the efficiency of the estimates required.

Data assimilation is ultimately an estimation problem and the methods that are generally applied originate from either estimation or control theory. Methods include sequential estimation using the Kalman filter, stochastic estimation using Maximum Likelihood, and direct minimisation methods such as genetic algorithms and simulated annealing. Details of these methods have been introduced in Chapter 2 but for more

specific information, the reader is referred to texts on data assimilation (e.g. Robinson, *et al.*, 1999; Brasseur, 1995).

6.2 MARINE ECOSYSTEM MODELLING

Scientists have attempted to model marine ecosystems since the late 1940's (e.g. Riley, 1946) and many models addressing different areas of marine ecosystems have been published. In the last two-three decades, there has been a significant revival of interest in these models as scientists wish to quantify the fluxes of the key growth limiting elements, carbon and nitrogen, at different scales, (Steele and Henderson, 1992). Models incorporating the increased knowledge of ecosystem components and the flows between them, have been recently developed (e.g. Pace *et al.*, 1984; Fasham 1985). These models are generally formulated to run over specific short time periods, such as over a spring bloom, and can be very sensitive to initial conditions. One further limitation of these models is their failure to produce realistic results when they are simulated for longer then the event timescale. As a result, a seasonal approach to modelling the nitrate-phytoplankton-zooplankton (NPZ) dynamics of the mixed layer was introduced by the seminal paper of Evans and Parslow (1985), where the annual cycle of the atmosphere and physical conditions in the sea drive the population cycle. Recent models have followed this approach (Matear, 1995; Fasham *et al.*, 1990). For a more detailed historical summary the reader is referred to Fasham, (1993) and references therein.

Although a stochastic approach to modelling ecosystems has been taken in the past (e.g. Jernigan and Tsokos, 1980; Kremer, 1983), the deterministic method is generally

favoured. The models are comprised of equations describing the dominant processes of the biological compartments, (deemed by the modeller to account for the main dynamics of the system), linked to a physical model governing the external environmental forcings of the system (e.g. climate variables and the depth of the mixed layer etc). This structure will tend to depend upon the motivation of the individual modeller, for example, according to whether the objective is to model individual species of zooplankton or zooplankton en masse.

The most recent models can be placed into two categories: firstly, complex models that strive to describe the complicated bio-geochemical processes within the ecosystem using coupled nonlinear differential equations, (e.g. Fasham *et al.*, 1990; Sarmiento *et al.*, 1993); secondly, simpler models that aim to expose the main dynamic features of the ecosystem and the effects of the external forcings on the system in a more qualitative manner (e.g. Evans and Parslow, 1985; Steele and Henderson, 1981). These deterministic models are highly parameterised: for example, the complex seven component model of Fasham *et al.*, (1990) and the simple model of Evans and Parslow (1985) have 26 and 15 parameters, respectively. Evidence in the recent literature suggests that there has been little critical evaluation of ecosystem model complexity in the marine sciences. Possible over-parameterisation and the resulting model uncertainty, derived from the difficulty in estimating parameters in over-parameterised models from limited data has not generally been investigated.

6.2.1 Selection of ecosystem model for analysis

Recent studies (Young *et al.*, 1996; Parkinson and Young, 1998) have demonstrated that complex nonlinear simulation models can be over-parameterised and that linearised, reduced order models can be identified that still retain the dominant dynamics of the system. For example, Parkinson and Young (1998) were able to identify a fourth order linear transfer function model that was capable of modelling atmospheric carbon dioxide to almost the same degree of accuracy as a 26th order nonlinear simulation model. It is possible, therefore, that the deterministic ecosystem model, outlined in Section 6.3, may be over-parameterised and can be reconstructed with reduced order. However, since these models are not particularly high order, major simplifications such as those achieved by the above researchers in connection with global climate models are unlikely.

Nevertheless, as a preliminary investigation, it is sensible to evaluate a simple lower order ecosystem model in order to determine whether over-parameterisation exists and model reduction is possible. In the present study, the three component model of Matear (1995) has been chosen for this kind of preliminary analysis. An assessment of the uncertainty in a relatively simple ecosystem model such as this, is essential, before investigating more complex models which it is hoped will form the basis of future research.

6.2.2 Mixed layer dynamics

The annual cycle of plankton populations are forced by the annual cycles of the physical conditions at sea: solar radiation; water temperature; nitrate concentration

below the mixed layer; day length; and the depth of the mixed layer. The most important controlling factors affecting plankton populations are the depth of the mixed layer and solar radiation. The mixed layer is defined as the water trapped above the thermocline. Below the mixed layer, the temperature of the water decreases rapidly and the thermocline describes the depth at which the greatest temperature change occurs. In the summer, the surface temperatures increase and the mixed layer depth reduces due to the development of a 'seasonal' thermocline. In the winter, there is a net loss of heat energy and perpetual wind mixing erodes away the seasonal thermocline; water temperatures are low and the mixed layer is deep. As the mixed layer deepens, nutrients are entrained from below, which fuels primary production; when the mixed layer reduces in the spring time, material is lost to the underlying layer. The phytoplankton that remain are subjected to increased daily irradiance which enhances their growth rate. It is traditionally assumed that rapid shallowing of the mixed layer causes a sudden increase in phytoplankton growth, resulting in the formation of a spring bloom (Gran and Baarud, 1935). An alternative theory is proposed by Evans and Parslow (1985), who, through the use of a deterministic model, found that sudden changes in the mixed layer depth were not the controlling factor triggering spring blooms. Their research suggests that the nature of the midwinter phytoplankton growth rate is the key to spring bloom development.

6.3 THREE COMPONENT NPZ ECOSYSTEM MODEL

The three component nitrate-phytoplankton-zooplankton, ecosystem model (Matear, 1995) (hereafter referred to as the TCE model) describes the exchange of nitrogen between three different biological compartments within the upper mixed layer of the

ocean. It describes the seasonal evolution of phytoplankton (P), whose growth rate is determined by photosynthesis and limited by essential nutrients (N). Phytoplankton are in turn grazed by herbivorous zooplankton (Z). The model assumes that the ocean is divided into two separate compartments. The mixed layer is biologically active, containing both plankton and nutrients with a uniform depth concentration. Nutrients, but no plankton are present in the biologically inactive layer below the mixed layer.

Following Evans and Parslow, (1985) the mixed layer dynamics of the ocean are not directly modelled. The rate of change of the observed mixed layer depth $\zeta(t)$

$$\frac{dM}{dt} = \zeta(t) \tag{6.1}$$

is used to modify the model in following asymmetrical manner. Deepening of the mixed layer $(\zeta(t) > 0)$ causes a change in concentration of non-motile components (components not capable of motion): water introduced from the inactive zone mixes with the surface water, diluting the concentration of phytoplankton, and causes mixing of surface and deep water nitrate. As no new water is introduced when the mixed layer shallows $(\zeta(t) < 0)$, the concentration of non-motile phytoplankton and nitrate remain constant and the concentration of motile zooplankton increases. This asymmetry is introduced into the model by defining $\zeta^{+}(t) = \max(\zeta(t), 0)$.

Following Matear, the coupled ordinary differential equations governing the temporal evolution of nitrate, phytoplankton and zooplankton concentrations take the following form:

$$\frac{dN}{dt} = -(\text{uptake} + \text{respiration}) + \text{mixing} + Z\ \text{excretion} + Z\ \text{predation} \tag{6.2}$$

$$\frac{dP}{dt} = (\text{uptake - respiration}) - \text{grazing by } Z - \text{sinking} \tag{6.3}$$

$$\frac{dZ}{dt} = \text{growth - mortality - sinking} \tag{6.4}$$

with the specific functional form:

$$\begin{aligned}\frac{dN}{dt} = &-[\alpha(t,M)Q_1 - \mu_1]P + \frac{m+\zeta^+(t)}{M}(N_o - N) \\ &+ (1-\gamma_2)\gamma_5 \frac{g(P-P_o)Z}{K_3 + P - P_o} + (1-\gamma_4)\gamma_5 Z\end{aligned} \tag{6.5}$$

$$\frac{dP}{dt} = [\alpha(t,M)Q_1 - \mu_1]P - \frac{g(P-P_o)Z}{K_3 + P - P_o} + \frac{m+\zeta^+(t)}{M}P \tag{6.6}$$

$$\frac{dZ}{dt} = \frac{\gamma_2 g(P-P_o)Z}{K_3 + P - P_o} - \mu_5 Z - \frac{\zeta(t)}{M}Z \tag{6.7}$$

where Q_1, the non dimensional nutrient limitation term is defined as,

$$Q_1 = \frac{N}{K_1 + N} \tag{6.8}$$

The photosynthetic growth rate of phytoplankton $\alpha(t,M)$ is averaged over the depth of the mixed layer and is obtained by integrating over one day/night cycle, i.e.,

$$\alpha(t, M) = \frac{2}{M} \int_0^{\tau} \int_0^{M} G(I)\, dz\, dt \tag{6.9}$$

The Smith (1936) function, $G(I)$, determines the phytoplankton growth rate as a function of light, and 2τ defines the day length. The descriptive parameters of the TCE model are defined with their *a priori* values in Table 6.1.

Table 6.1 Ecosystem model parameters and suggested *a priori* values (Matear, 1995)

Definition	Symbol	A *priori* value	Units
Phytoplankton Parameters			
Half Saturation Constant of N uptake	K_1	1.0±1.0	μM
Specific Mortality Rate	μ_1	0.024±0.05	d^{-1}
Light attenuation by P	K_c	0.06	$m^{-1}(\mu M)^{-1}$
Light attenuation by water	K_w	0.04	m^{-1}
Initial slope of the P-I curve	α	0.025±0.01	$(Wm^{-2}d)^{-1}$
Phytosynthetically active radiation	PAR	0.5	-
Zooplankton Parameters			
Assimilation efficiency of Z	γ_2	0.5±0.1	-
Detrial fraction of Z mortality	γ_4	0.33±0.1	-
Fraction of Z excreted as N metabolites	γ_5	0.6±0.05	-
Specific mortality of Z	μ_5	0.07±0.02	d^{-1}
Maximum growth rate of Z	g	1.0±0.2	d^{-1}
Half saturation for ingestion of Z	K_3	0.5±0.3	μM
Grazing threshold of Z	Po	0.05±0.05	μM
Physical Mixing Parameters			
Diffusion Rate	m	3.0±0.5	md^{-1}

6.3.1 SIMULINK representation of the ecosystem model

Previously, ecosystem models have mainly been simulated after extensive programming in FORTRAN or other similar computer languages (Fasham *et al.*, 1990). In this study, a more novel approach has been used and the ecosystem model has been constructed using SIMULINK, an 'iconographic' dynamic simulation computer package and an extension of the numerical and visualisation package MATLAB. Complex nonlinear models are easily simulated utilising SIMULINK: using the iconographic libraries, linear and nonlinear icon blocks can be selected and connected together graphically using mouse driven commands to quickly construct a model. Utilising the whole block diagram, the model equations are collated and solved using a user-specified integration algorithm. This novel approach has the advantage that models can be formulated and analysed using a single software environment. Other key benefits of using SIMULINK comprise: an enhanced visualisation of the model; greater interaction with the model allowing adjustments to be made easily as the research dictates; change of parameter values during simulation; and lastly the availability of output data to MATLAB for extended numerical analysis. The top layer of the SIMULINK iconographic ecosystem model is shown in Figure 6.1. Here the blocks 'PHYTOPLANKTON','NITRATE', and 'ZOOPLANKTON' contain the model equations in block form and these are shown in Figures 6.2-6.8.

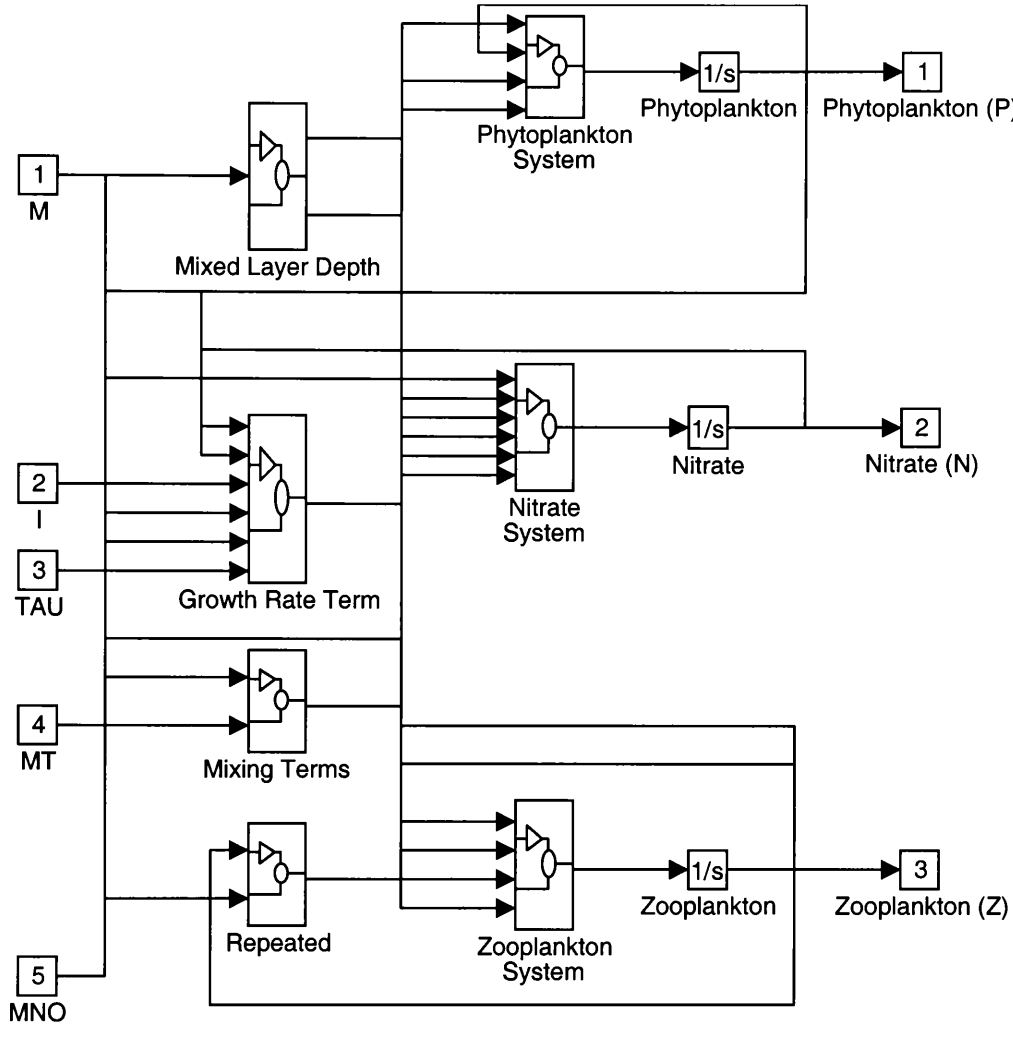

Figure 6.1 SIMULINK iconographic representation of the TCE Model

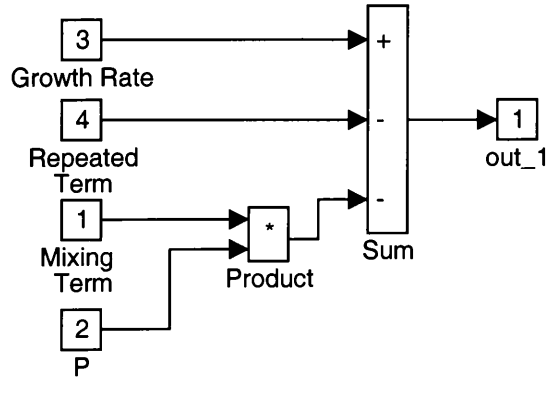

Figure 6.2 SIMULINK iconographic representation of the Phytoplankton System

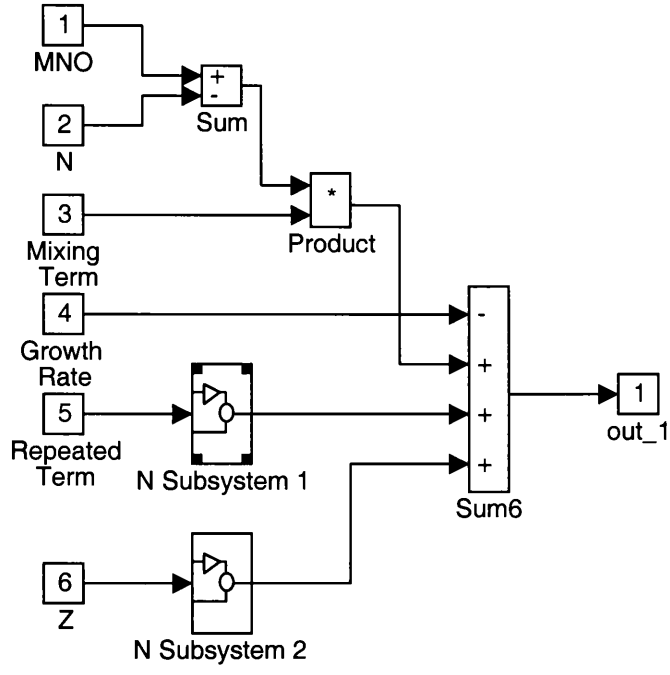

Figure 6.3 SIMULINK iconographic representation of the Nutrient System

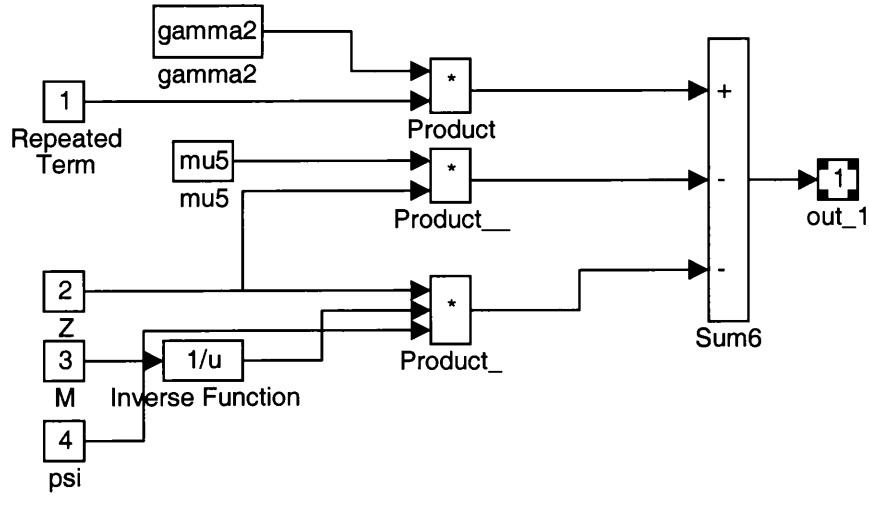

Figure 6.4 SIMULINK iconographic representation of the Zooplankton System

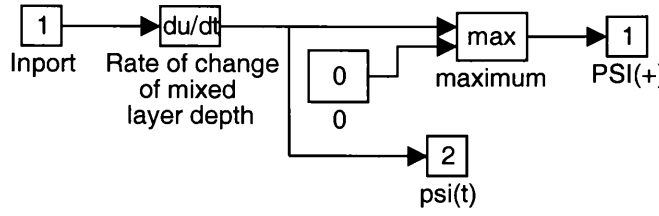

Figure 6.5 SIMULINK iconographic representation of the mixed layer depth equations.

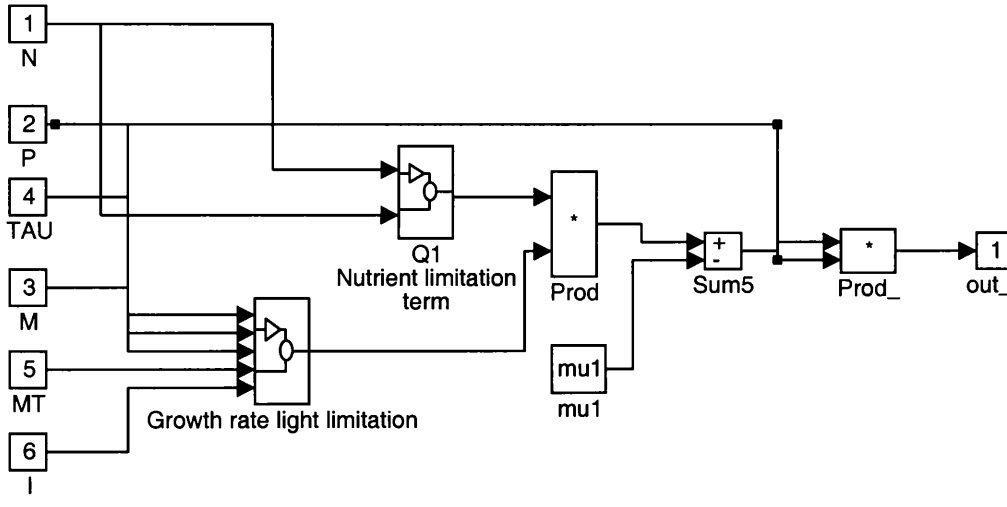

Figure 6.6 SIMULINK iconographic representation of the growth rate term.

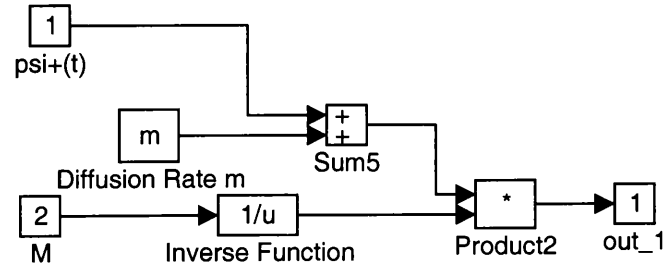

Figure 6.7 SIMULINK iconographic representation of the mixing terms.

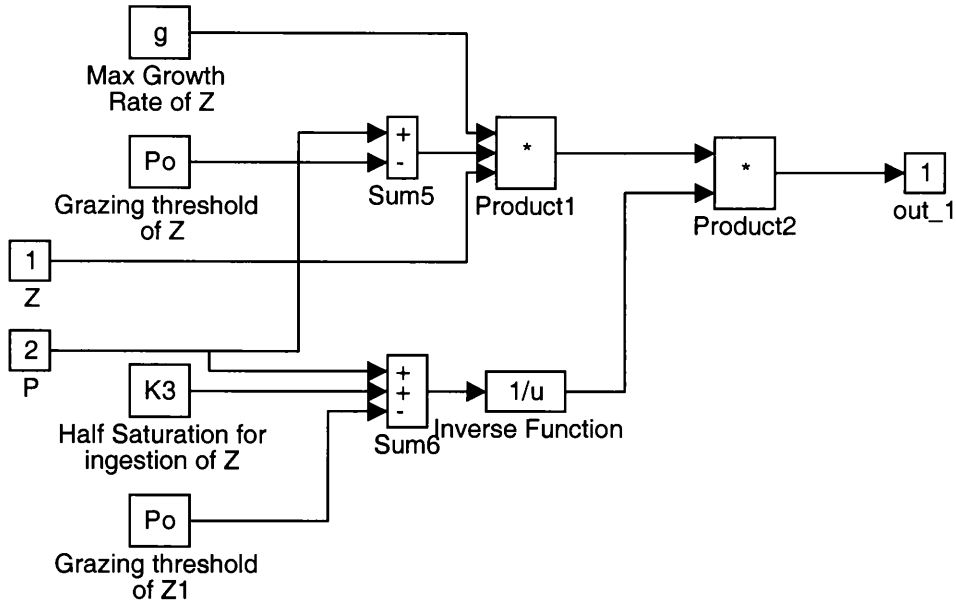

Figure 6.8 Repeated term.

6.3.2 Biological data.

The data used for this analysis originate from observations taken at the Ocean Weather Station P (50°N, 145°W), in the Gulf of Alaska. The data can be separated into three categories. Firstly, the data that are required to drive/force the model dynamics. Secondly, observations of the model outputs used to constrain the model parameters during optimisation; and thirdly, suggested a *priori* values for the model parameters.

The ecosystem model is driven by five forcing terms that show characteristic annual cycles: mixed layer depth, solar radiation, day length, temperature of the mixed layer and nutrient concentration immediately below the mixed layer and are shown in Figure 6.9. These data are based on average measurements taken at Station P.

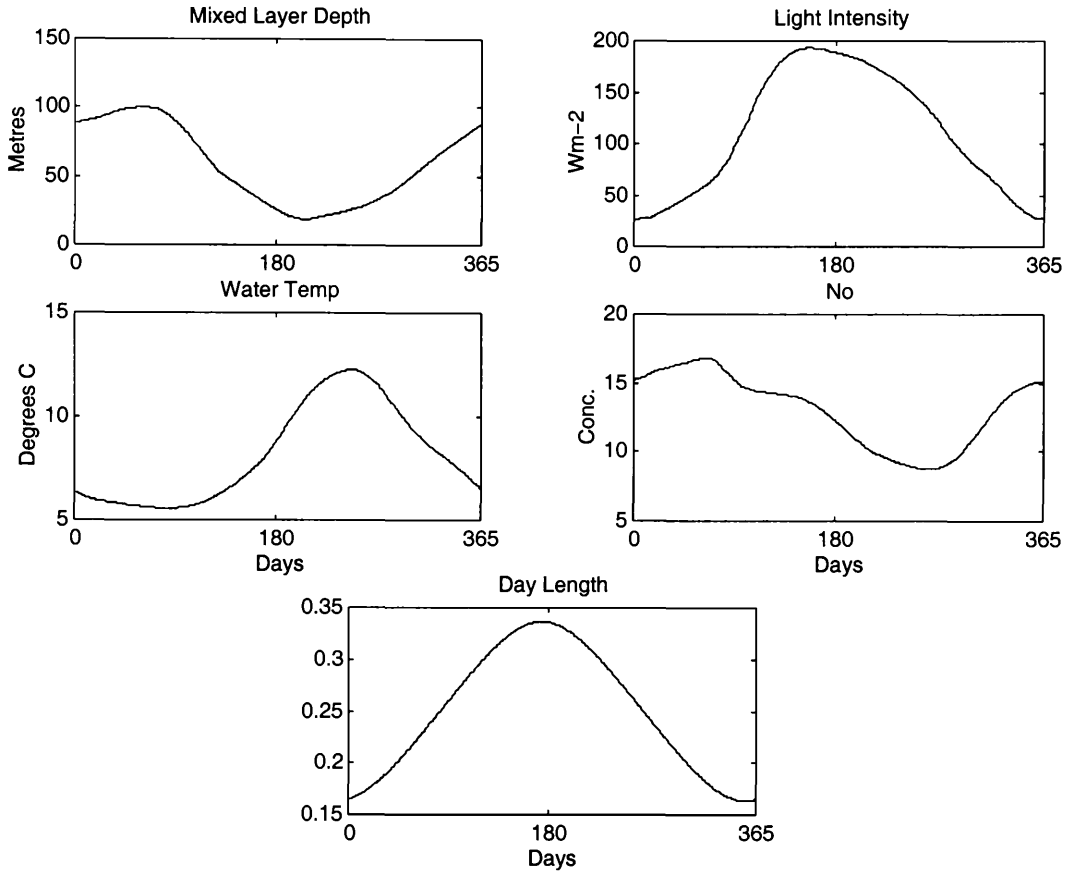

Figure 6.9 Physical forcing terms of the ecosystem model

Nitrate (N), phytoplankton (P), net phytoplankton productivity (NPP) and limited zooplankton (Z) observational data, taken at Station P, are available to optimise the model parameters, as shown in Figure 6.10. NPP is defined as the net uptake of CO_2 by phytoplankton for photosynthesis, less the CO_2 respired. The phytoplankton concentration at Station P has little seasonal variation, which is common throughout the open subarctic Pacific (Frost, 1991). This contrasts with the seasonal variations at shelf and inland subarctic Pacific waters and in other oceans (e.g. sub-arctic north Atlantic) where pronounced seasonal blooms are observed (see e.g. Landry *et al.*, 1989; Kanda *et al.*, 1990; Parson and Lalli, 1988). The phytoplankton stock remain reasonably constant despite the strong seasonal variations in the properties of physical

forcings that significantly affect the growth rate as presented in Figure 6.10. Different theories accounting for this atypical behaviour are discussed by Frost (1991). From Figure 6.10, we see that the nitrate concentration is maintained throughout most of the summer months whereas in the entire open sub-arctic Pacific, nitrate is depleted. A summer peak of zooplankton biomass is observed.

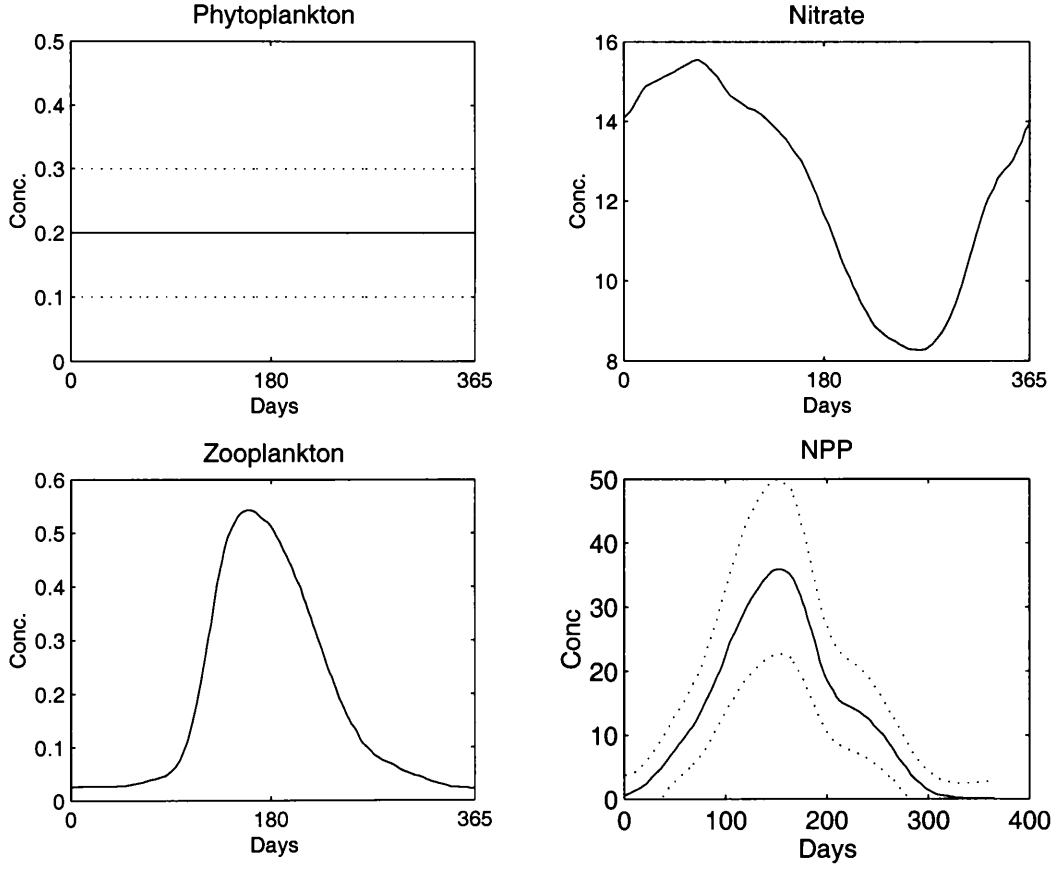

Figure 6.10 Smoothed observed nitrate, phytoplankton, NPP and zooplankton concentration with ± one standard deviation uncertainty

Full detailed descriptions of all the data and their relative uncertainty can be found in Matear (1995) and references therein. For convenience, the data used in this study have been carefully digitised from figures featured in Matear (1995).

6.3.3 Data for parameters and calibration

Assigning values to the parameters that specify the ecological interactions in a model can be difficult, especially as biological parameters, unlike most physical and chemical parameters, cannot be strictly considered as constant. Whilst some parameters can be measured experimentally at sea, others are either impossible to measure or they can not be measured accurately. Traditionally, one approach around this problem has been to assign values to those parameters, where experimentally determined data are available and to simply adjust the remaining parameters until a good fit to the data is achieved. Following this *ad hoc* constrained optimisation procedure, it is often difficult to obtain a good fit to the observation data and to determine whether this is due to inadequacies in the model structure, or whether the appropriate parameter space has been considered.

6.4 REVIEW OF OPTIMISATION METHODS

Three different nonlinear optimisation techniques applied to oceanic ecosystem models have been presented recently in the literature: conjugate direction, Markov Chain Monte Carlo (MCMC) and simulated annealing methods are outlined by Fasham and Evans, (1995), Harmon and Challenor, (1997) and Matear, (1995) respectively. Several key points can be deduced from the applications of these optimisation methods.

Firstly, it is not always possible to obtain unique parameter sets that give an optimal fit to the data (Matear, 1995; Fasham and Evans, 1995). Despite (or perhaps because of) the complexity of the ecosystem model and the large number of parameters required, Fasham and Evans (1995) were unable to find a unique parameter set that was able to

give a good fit to all the data simultaneously. Two different optimal sets of parameters were discovered, one gave a good fit to primary production but a bad fit to zooplankton; and the other produced the converse results. Secondly, high parameter cross correlation was often observed at the optimisation termination, suggesting possible over-parameterisation of the model. Simulating the results of parameter estimation using simulated annealing (Matear; 1995) revealed that for several different models, only 10 independent parameters could be estimated from the data. Thirdly, the Markov Chain Monte Carlo (MCMC) method adopted by Harmon and Challoner (1997) also exemplified this point, as it became impractical to estimate more than 10 parameters from a total of 28.

6.5 DETERMINISTIC PARAMETER OPTIMISATION

Initial simulation of the TCE model with different parameter sets indicated how sensitive the model was both to initial conditions of the three states (*N, P, Z*) and to the parameter values. Inappropriate choice of the parameter values caused the model either to go highly unstable or yield unrealistic results. Furthermore, the quasi-Newton and nonlinear least squares optimisation routines, FMINS, FMINU and LEASTSQ (MathsWorks, 1992), were unable to converge unless they were initialised with parameter values close to those at the global minima in the hyper-space. In contrast, initiation problems were not encountered through optimising the model parameters using a genetic algorithm (GA), a stochastic global search method (see Chapter 2). GA's utilise a 'population' of potential solutions (parameter sets) to avoid getting stuck in local, rather than the global minima in the parameter hyper-space. For these reasons, a two stage process of optimisation was carried out. Firstly, the genetic

algorithm was used to locate the approximate area within the parameter space which contained the true minima. Secondly, the parameter estimates were refined using a nonlinear least squares algorithm which additionally generates a parameter covariance matrix, providing estimates of the precision and the sensitivity of the model to each parameter.

Model parameter optimisation was carried out twice using two different multi-objective cost functions. The first, equation (6.10) was simply used to minimise the data misfit only. The second, equation (6.11) included data misfit, steady state and *a priori* parameter penalty terms. Using two separate cost functions in the optimisation analysis allowed, firstly, to obtain the set of parameters that gave the best model fit to the data and secondly, the set of parameter values that gave both the best model fit when restricted by initial *a priori* parameter limits. The model was simulated for three years during each optimisation iteration so that the model had reached a steady operating state, and the output from the final year was utilised within the cost function. For simplicity, parameters K_w, K_c and PAR were held constant during each optimisation iteration,

$$J = \frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_i - \hat{\mathbf{y}}_i)^2 \tag{6.10}$$

$$J = \frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_i - \hat{\mathbf{y}}_i)^2 + \frac{1}{2}\sum(\mathbf{y}_1 - \hat{\mathbf{y}}_{366})^2 + \frac{1}{2}\sum_{i=1}^{m}(\mathbf{z} - \bar{\mathbf{z}})^2 \tag{6.11}$$

where matrix $\mathbf{y}$ defines the observed concentrations of N, P, Z and NPP $\left(\mathbf{y} = [N_i P_i NPP_i Z_i]\right)$; and the matrix $\hat{\mathbf{y}}$ defines the simulated outputs at day i of the

total n days of the simulation period. The second term in equation (6.11) forces the model to have a steady state output over the seasonal cycle. The final term in equation (6.11) is incorporated to force the m model parameters $\mathbf{z}$ to maintain a close approximation to their expected *a priori* values $\bar{\mathbf{z}}$.

Repeated optimisation highlighted the complex nature of the parameter hyperspace and even with the inclusion of the GA, local minima were not always avoided. Each separate iteration converged on completely different sets of parameters, with only slightly different values of the cost function. Many successive iterations of the optimisation procedure were required, therefore, to ensure that the full parameter hyperspace was sampled. Maximum and minimum values for each parameter were user-specified within the GA routine. These boundaries were often reached as the GA attempted to converge, especially when the first cost function (6.10) was used. The resulting parameter estimates in such cases were then already biased. In light of this evidence, the parameter value boundaries were relaxed to allow the GA to move in an unrestricted manner around the parameter space. However, regardless of the boundary width they were often reached. These results could suggest parameter collinearity associated with possible model over-parameterisation.

The parameter sets that minimised the two cost functions are shown in Table 6.2 as Set 1 and 2 respectively. The outputs from simulating the model using these parameter sets are compared to the observational data in Figure 6.11.

Table 6.2 Optimal model parameters with standard errors given in parenthesis.

Parameter	Set 1	Set 2
K_1	3.0001 (4.0651)	0.1385 (1.6623)
μ_1	0.0034(0.0043)	0.0006 (0.0127)
Po	0.1835 (0.0435)	0.0830 (0.0219)
K_3	1.9234 (20.5315)	0.4388 (2.4308)
g	1.9309 (19.4843)	1.9962 (9.2978)
γ_2	0.3717 (0.0806)	0.5809 (0.7005)
μ_5	0.0074 (0.0016)	0.0630 (0.0190)
m	1.5308 (0.4489)	2.1875 (0.7180)
γ_5	0.0001 (1.1608)	0.5858 (12.9320)
γ_4	0.0007 (2.7910)	0.3210 (9.3211)
α	0.0219 (0.0075)	0.0500 (0.0159)
Cost Function 1	0.264	0.686
Cost Function 2	7.170	0.4131

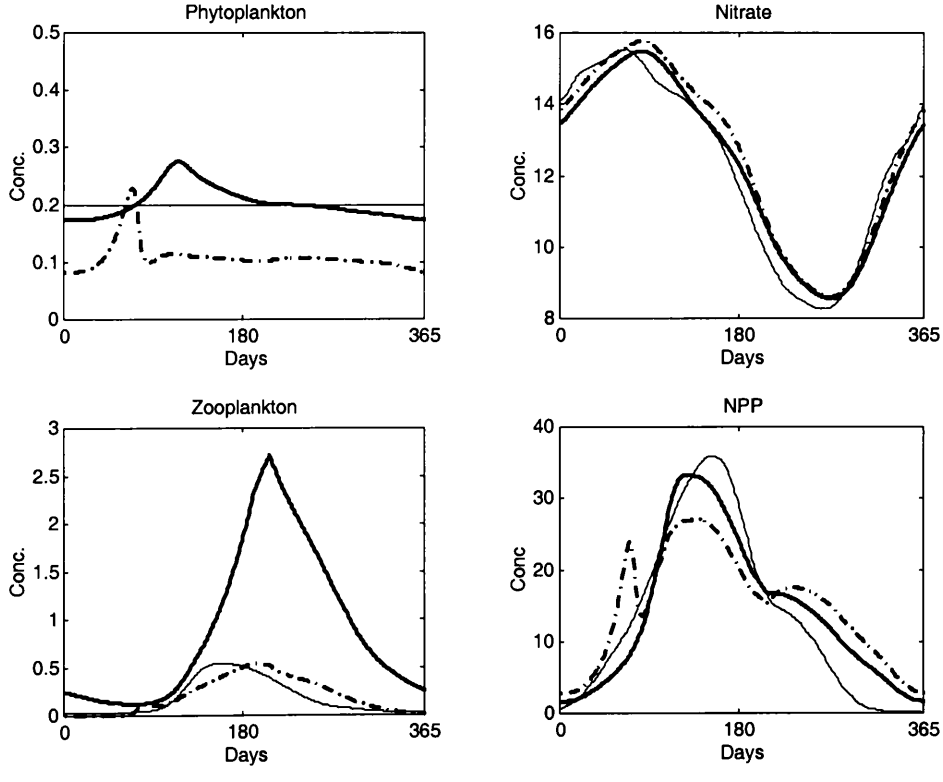

Figure 6.11 Observation data (thin line) and simulation response to optimised parameter set 1 (bold line) and 2 (dotted line). Units of concentrations are $\mu M\, N\, m^3$

Parameter Set 1 successfully generates a sustained limit cycle for all outputs and reproduces the observed phytoplankton, nitrate, and NPP concentrations well within the uncertainty limits (one standard deviation) of the data (Figure 6.10). However, the simulated zooplankton concentration is much higher than the observed data, a problem that was consistently observed throughout most optimisation iterations. The failure of the simulation to capture the zooplankton dynamics, highlights either a limitation of the model structure and/or inappropriate model parameter values. It should be noted, however, that the measured zooplankton data series has the greatest uncertainty and this may account for these problems.

The observed nitrate, *NPP* and zooplankton concentrations are successfully reproduced using the parameter Set 2. The close model fit to the observed zooplankton concentration is almost unique to these parameter values: no other set of parameters from all the optimisation solutions showed this behaviour. However, the phytoplankton concentration is significantly lower in comparison to the concentration obtained with parameter Set 1. The model generates a weak spring phytoplankton bloom but, for the majority of the remaining annual cycle, the phytoplankton concentration drops just below the uncertainty limits of the data. The spring bloom is the result of a disparity between the growth of phytoplankton and their grazing by zooplankton. This, in turn, is caused because of the existence of a zooplankton feeding threshold within the model. Subsequent to the phytoplankton concentration exceeding this threshold, grazing by zooplankton reduces the phytoplankton concentration to a steady level.

The above results are not surprising as Matear (1995) was unable to obtain a parameter set that gave a model output that simultaneously fitted all four observed series from Station P, without imposing an artificial constraint on the model. Consequently, the ability to capture the dynamics of the annual zooplankton concentration with only a small effect on the phytoplankton, by using parameter Set 2 can be regarded as an achievement of the present optimisation.

The best model fit to the data (in cost function terms) was produced using parameter Set 1, which was obtained giving the optimisation algorithms unrestricted access to the parameter space. However, a large disparity exists between the optimised and *a priori* parameter values; only three parameters, μ_1, γ_2, and α have values within the '*a priori*' range. Note also that the parameter penalty component, incorporated within cost function 2 did not restrict parameters α, m, and g from Set 1 to be within the *a priori* range.

If the optimised values are so unrealistic in physical terms the whole parameter set could be completely rejected. In this manner, the model structure could be regarded as an unrealistic representation of the ecosystem at Station P. The outcome of such a debate will depend largely upon the goals of the modelling exercise and the perception of the modeller i.e. whether to produce a model that will achieve the best fit to the data, or to produce a model that must retain the greatest relationship to physical reality as currently perceived, so compromising model fit.

It is important to consider the high standard errors, associated with each parameter estimate, presented in Table 6.2. Both models are badly defined; only five out of the

eleven parameters in Set 1 and three parameters in Set 2 have standard errors smaller in magnitude than their respective parameter value. The covariance matrix of each parameter set can be utilised to investigate this uncertainty. The covariance and associated correlation matrices indicate that a number of parameters are closely correlated. Table 6.3 presents the correlation matrix of parameter Set 1. Six pairs of parameters have an absolute correlation coefficient greater than 0.75 (highlighted in bold), indicating that only five parameters out of eleven can be independently identified. The results from repeated optimisation iterations, indicated that three pairs of parameters, K_3:g, γ_4:γ_5 and α:μ_1, consistently showed the highest correlation. The correlation matrix of parameter Set 2 exhibits similar characteristics.

Table 6.3 Correlation coefficients of the parameters in Set 1

	K_1	μ_1	Po	K_3	g	γ_2	μ_5	m	γ_5	γ_4	α
K_1	**1.000**	0.736	0.745	**-0.757**	**-0.753**	-0.312	-0.729	0.055	0.009	-0.050	**0.751**
μ_1	0.736	**1.000**	0.225	-0.635	-0.627	-0.429	-0.650	0.056	0.058	0.017	**0.931**
Po	0.745	0.225	**1.000**	-0.650	-0.650	0.048	-0.376	0.047	-0.019	-0.030	0.141
K_3	-0.757	-0.635	-0.650	**1.000**	**1.000**	0.280	0.597	-0.010	-0.068	-0.012	-0.582
g	-0.753	-0.627	-0.650	**1.000**	**1.000**	0.267	0.594	-0.004	-0.063	-0.008	-0.575
γ_2	-0.312	-0.429	0.048	0.280	0.267	**1.000**	0.443	0.139	0.086	0.180	-0.502
μ_5	-0.729	-0.650	-0.376	0.597	0.594	0.443	**1.000**	-0.450	-0.163	-0.183	-0.741
m	0.055	0.056	0.047	-0.010	-0.004	0.139	-0.450	**1.000**	0.490	0.627	0.109
γ_5	0.009	0.058	-0.019	-0.068	-0.063	0.086	-0.163	0.490	**1.000**	**0.836**	0.086
γ_4	-0.050	0.017	-0.030	-0.012	-0.008	0.180	-0.183	0.627	0.836	**1.000**	0.032
α	0.751	0.931	0.141	-0.582	-0.575	-0.502	-0.741	0.109	0.086	0.032	**1.000**

Matear (1995) also observed parameter cross correlation of this type when optimising the TCE model using simulated annealing. However, the correlated parameter pairs (K_1:μ_1, K_1:θ, K_1:α, μ_1:θ, μ_1:α, θ:α) he identified are not all identical to those identified above. In particular, the optimisation results above and Matear's indicate different parameter pairs exhibit cross-correlation in different regions of the parameter

hyper-space. This unusual behaviour may be a direct result of the model nonlinearity operating in a contrasting manner depending on the region of the parameter hyper space. To illustrate parameter collinearity, an optimisation iteration was performed where only K_3 and g were free to vary. As expected, the optimisation algorithms were unable to converge and a trace of the two parameter values during the optimisation showed the parameter values changing in order to maintain a fixed ratio with each other. It is likely that multi-parameter correlation exists, which cannot be so easily evaluated.

Table 6.4. Re-optimisation of Solution 1 parameters. Parameters marked with an asterix are held constant.

Parameter	Value	Standard Error
K_1	3.0018	2.7092
μ_1*	0.0034	-
Po	0.2189	0.0456
K_3*	1.9234	-
g	1.9296	0.4385
γ_2	0.3490	0.0770
μ_5	0.0066	0.0013
m	1.5305	0.3961
γ_5	0.0001	0.7070
γ_4*	0.0007	-
α	0.0177	0.0025

It is clear that uncertainty can be propagated through the model as a result of this high parameter correlation. Table 6.4 presents the results of holding one parameter from each of the three identified pairs of highly correlated parameters constant (K_3:g,

γ_4:γ_5 and α:μ_1). The parameter uncertainty has been reduced: the standard errors of the 8 estimated parameters are much lower than their original values (see Table 6.2).

Having defined the correlated parameters, examination of the model structure clearly shows there is indeed a direct relationship between the effects of the correlated parameters (Equation 6.5-6.7). For example, phytoplankton mortality rate μ_1 and α, the initial slope of the P-I curve, both fundamentally control the growth rate of phytoplankton and consequently a change in μ_1 can be compensated by α to maintain an output level and *vice versa.* Again parameters g and K_3 (maximum growth rate and half saturation for ingestion of zooplankton, respectively) both determine the growth of zooplankton. Mathematically, the two parameters form a ratio which makes their correlation understandable.

6.6 MODEL UNCERTAINTY AND SENSITIVITY ANALYSIS.

All environmental models are approximations of reality and are, therefore, subject to uncertainty at many different levels. A broad distinction between the sources of uncertainty can be made (Beck, 1993). Firstly, uncertainty arising from the internal description of the model: spatial, temporal and ecological aggregation; model structure (Beck, 1985); a priori values of model parameters and states; and numerical errors in the solution. Secondly, uncertainty arising from the external description of the system: measurement errors associated with system inputs and outputs; and unobserved system disturbances. Many formal procedures of assessing the effects of uncertainty have been proposed (see e.g. Beck, 1987). However, the qualitative assessment of all uncertainty sources presents a large temporal burden. Model structure or the

deterministic equations governing the model, are usually viewed as hypotheses which can be revised in light of evidence derived from other uncertainty and sensitivity analysis as discussed by Kremer, (1983).

Sensitivity analysis is used to assess and quantify the effects of model uncertainty and methods of analyses can be broadly classified into two main categories.

1. Deterministic: Classical *analytical* analysis of the model.
2. Stochastic: Analysing the model in *probabilistic* terms.

Both methods have the overriding aim of quantifying the impact of changes to the different model components on the overall model output; and large impacts imply that the component is important in controlling model behaviour. Traditional analytical methods (see review by Thornton, 1993; Tomovic, 1963) such as using partial derivatives of the output with respect to the parameters have been used extensively. However, models that are complex and incorporate nonlinearities are more difficult to study in analytic terms, particularly where there is a possibility of strong parameter interactions. Stochastic approaches to sensitivity analyses (e.g. Kremer, 1983; Rose, 1993; Hornberger and Spear, 1980; Young *et al.,* 1996) are more powerful and have been applied to wide range of disciplines.

The uncertainty of lake ecosystem models has been evaluated quite extensively (Fedra *et al.*, 1981, Scavia, 1981; 1993; van Straten, 1980; 1983) but only limited parameter sensitivity analysis of ecosystem models has been reported in the literature (Fasham *et al.*, 1990 and Jambart; 1979). However, their research evaluated the sensitivity of a

model to each parameter individually, disregarding the effects of parameter interactions.

The sensitivity analysis approach adopted in the present study is based on Monte Carlo simulation (MCS) and associated Generalised Sensitivity Analysis (GSA) (Spear and Hornberger, 1980), which comprehensively accounts for parameter interdependence.

6.6.1 Stochastic simulation modelling

In general, a deterministic nonlinear simulation model can be defined in probabilistic terms by a set of lumped parameter, vector differential equations taking the following form:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(x(t), \mathbf{u}(t), \zeta(t), \vartheta) \tag{6.12}$$

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\xi}(t), \boldsymbol{\vartheta}) \tag{6.13}$$

in which $\mathbf{x}(t) = [x_1\ x_2 \ldots x_m]^{\mathrm{T}}$ is an m-dimensional vector of state variables that characterise the system; $\mathbf{u}(t) = [u_1\ u_2 \ldots u_n]^{\mathrm{T}}$ is an n-dimensional vector of measured inputs or exogenous disturbances to the system; $\boldsymbol{\zeta}(t) = [\zeta_1\ \zeta_2 \ldots \zeta_p]^{\mathrm{T}}$ is a p-dimensional vector of unobserved stochastic inputs (the system *noise*); $\boldsymbol{\vartheta} = [\vartheta_1\ \vartheta_2 \ldots \vartheta_q]^{T}$ is an q-dimensional vector of model parameters within the defined nonlinear structure $\mathbf{f}$; $\mathbf{y}(t) = [y_1\ y_2 \ldots y_r]^{\mathrm{T}}$ is an r-dimensional vector of observations

of the system output; and $\boldsymbol{\xi}(t) = [\xi_1\ \xi_2 \ldots \xi_s]^T$ is an s-dimensional vector of observational or measurement noise, defined in stochastic terms.

Utilising the model formulated by equations (6.12-6.13) two sources of uncertainty can be defined: uncertainties in the measured and stochastic inputs, $\mathbf{u}(t), \boldsymbol{\zeta}(t)$, and uncertainty in the parameters, $\boldsymbol{\vartheta}$. In the probabilistic uncertainty and sensitivity analysis, an ensemble of model outputs, generated through repeated MCS, can be assessed in order to examine and quantify the propagation of uncertainly through the model. For each random MCS realisation, each parameter is defined by randomly sampling from its predefined parent probability density function (pdf). The model is simulated using these constant parameters, perturbed by either stochastic or deterministic inputs (or both).

6.6.2 Generalised sensitivity analysis

Generalised Sensitivity Analysis (GSA), an extension of MCS, is an objective procedure for evaluating the relative significance of a parameter and its effect on a specific mode of model behaviour over a simulation period. GSA was first developed by Spear and Hornberger, (Spear and Hornberger, 1980; Hornberger and Spear, 1980) and has had many applications, including different environmental investigations e.g. water quality, (Whitehead and Young, 1979; Whitehead and Hornberger, 1984; Beck, 1987; Jakeman *et al*, 1990b), climate (Parkinson and Young, 1998), ecology (Scavia, 1993), control engineering (Spear, 1970) and micro economics (Young *et. al*, 1973).

Prior to the GSA analysis, a range of responses for each output of the model, deemed to be representative and unrepresentative of the observed system behaviour, are defined. Repeated simulations, using randomly selected parameter sets from predefined pdf's are then performed and each randomly chosen parameter set is classified as either producing a behavioural (B) or non-behavioural model response $(\overline{B})$. For each model parameter the Kolmogorov Smirnov (KS) two sample test can be used to identify whether there is any significant statistical difference between the values in the B and $\overline{B}$ sets obtained in the MCS analysis. The cumulative probabilistic distribution of each parameter from both n behavioural and m non-behavioural sets can be formed and the maximum vertical distance $d_{m,n}$ between them derived. The distance $d_{m,n}$ can be subsequently checked with the KS statistic to confirm its relative significance.

The covariance matrix, derived from the original optimisation shown in Table 6.3, can be used as an objective tool for constructing the pdf for each parameter. The parameter values for each stochastic realisation are drawn at random from these pdf's, so that the collinear or interdependent relationship between parameters is maintained in the MCS analysis. Note, however, that the differentiation between *behavioural* and *non behavioural* class is extremely subjective and this is exacerbated in this case by the existence of the multiple outputs from the model.

6.6.3 Uncertainty Analysis of the TCE Model using GSA

For this study parametric uncertainty was considered paramount and, therefore, model inputs were considered as either measured or deterministic rather than stochastic. Following the MCS/GSA procedure, the main aims of the GSA analysis were to:

1. examine how uncertainty in the parameters is propagated through the model;
2. identify which parameters were the most significant in affecting model responses;
3. compare results of (2) in light of conclusions drawn from the initial model optimisation and discussed in Section 6.5;
4. on the basis of (2) and (3), identify the subset of significant model parameters for future optimisation research.

For the purpose of the present GSA, a simulation was classed as $\overline{B}$ if any of the model outputs had a concentration that was greater or less than the observed series plus/minus 1.5 standard deviations. However, classification was limited to the phytoplankton, NPP and nitrate concentration outputs, since the uncertainty associated with the zooplankton data and the model's failure to reproduce these data make it impossible to consider the zooplankton sensibly.

The initial model optimisation results discussed in Section 6.5 have shown that most parameters were badly estimated as a direct result of inter-parameter correlation. Consequently, pdf's formulated have very wide distributions. Randomly selecting parameter values from such pdf's inevitably leads to values being chosen with large deviations from their respective means. As a result, it might be assumed that the model would be sensitive to such shifts and produce an output unrepresentative of the system. However, by virtue of the parameter collinearity retained by the selection

procedure, the pair ratio value remains approximately consistent and the model output is still representative of the system. As a result, it proved necessary to scale the covariance matrix to extend the range of each parameter pdf, in order to allow a proportionate number of B and $\overline{B}$ simulations thus enabling proper use of the KS statistics. In particular, the model was separately simulated with parameters randomly sampled from pdf's defined from the mean parameter Sets 1 and 2 and their covariance matrices, scaled by 10 and 1, respectively. Using this approach the pdf's of some parameters extended below zero, which would give parameters with values that are physically speaking unfeasible. In these circumstances, whereever a randomly selected set of parameters contained a value below zero the whole set was disregarded. Consequently, a skewed rather than a Gaussian parameter distribution was randomly sampled, which could, therefore, be regarded as a limitation of the technique. Typical B and $\overline{B}$ responses for both parameter sets are shown in Figure 6.12.

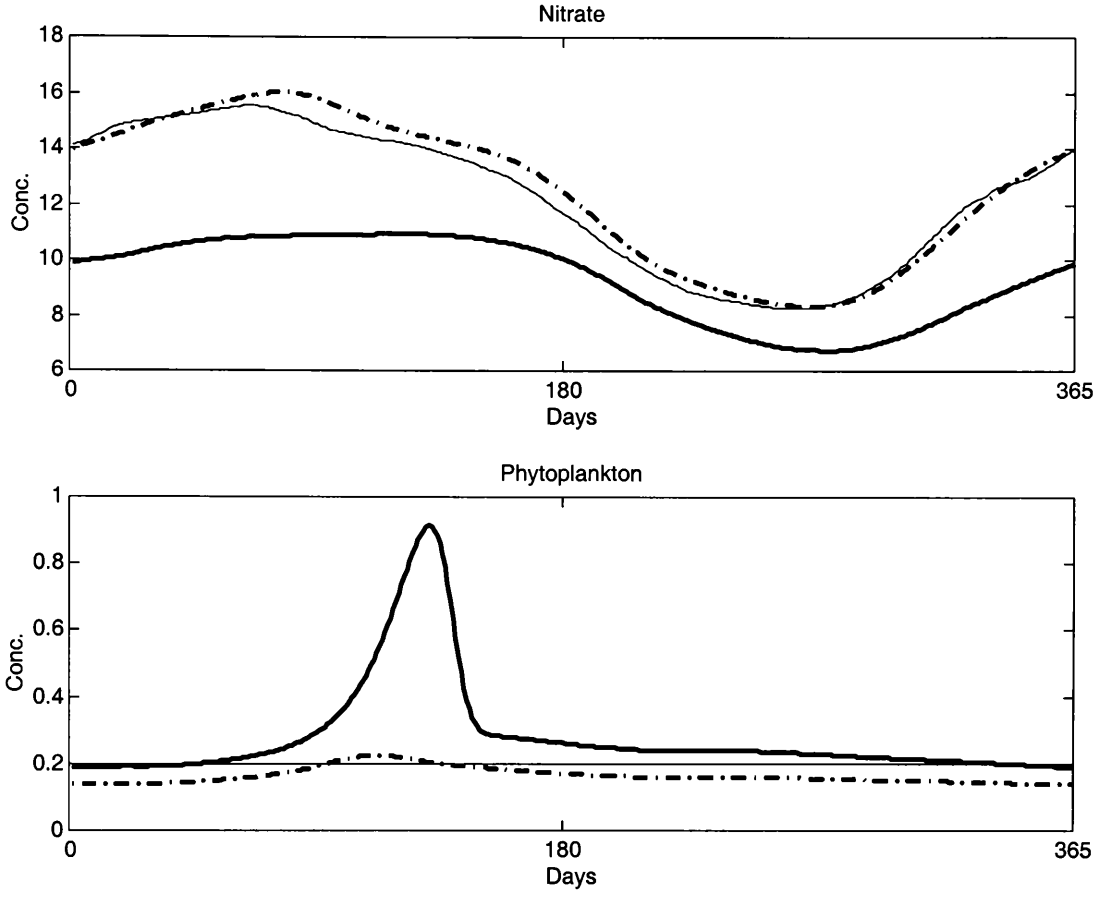

Figure 6.12 Typical examples of a behavioural (dotted line) and non-behavioural (bold line) model response for nitrate and phytoplankton. Observed data (thin line).

The results of the GSA analysis for both parameter sets are shown in Table 6.5. The value of the KS statistic that defines 95% statistical significance is 0.0451 and 0.0493 for Sets 1 and 2 respectively. KS statistics give a value between 0-1, the higher the value the more significant the parameter. Insignificant parameters of the TCE model are highlighted by an asterix. However, as we see, all the parameters show statistical significance with the exception of K_1 from Set 2. This implies that all the parameters are required by the model, in order to generate a behavioural response.

Table 6.5 Significant parameters of the TCE model determined by GSA.

Parameter	Set 1: K-S Result	Set 2: K-S Result
K_1	0.1300	0.0472*
μ_1	0.1057	0.1050
Po	0.2799	0.1129
K_3	0.0947	0.1732
g	0.1101	0.0735
γ_2	0.0881	0.4383
μ_5	0.1872	0.0525
m	0.2225	0.1102
γ_5	0.2555	0.3333
γ_4	0.0969	0.2595
α	0.1410	0.1207

For illustration, cumulative frequency distributions (cfd's) for parameters m (Set 1) and K_1 (Set 2), are presented in Figures 6.13 and 6.14 respectively. The horizontal dashed line signifies the KS statistic. The dotted line indicates the magnitude of difference between the two responses. Figure 6.13 shows behavioural (B) and non-behavioural ($\overline{B}$) responses that are 'significantly different'. Figure 6.14 shows behavioural B and $\overline{B}$ responses that are 'not significantly different'.

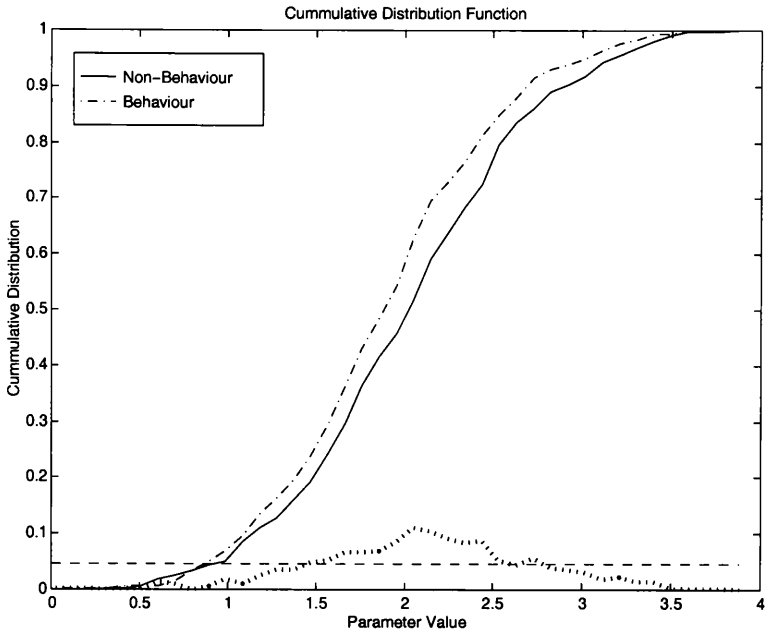

Figure 6.13 Significantly different B and $\overline{B}$ responses observed for parameter Set 1 parameter m

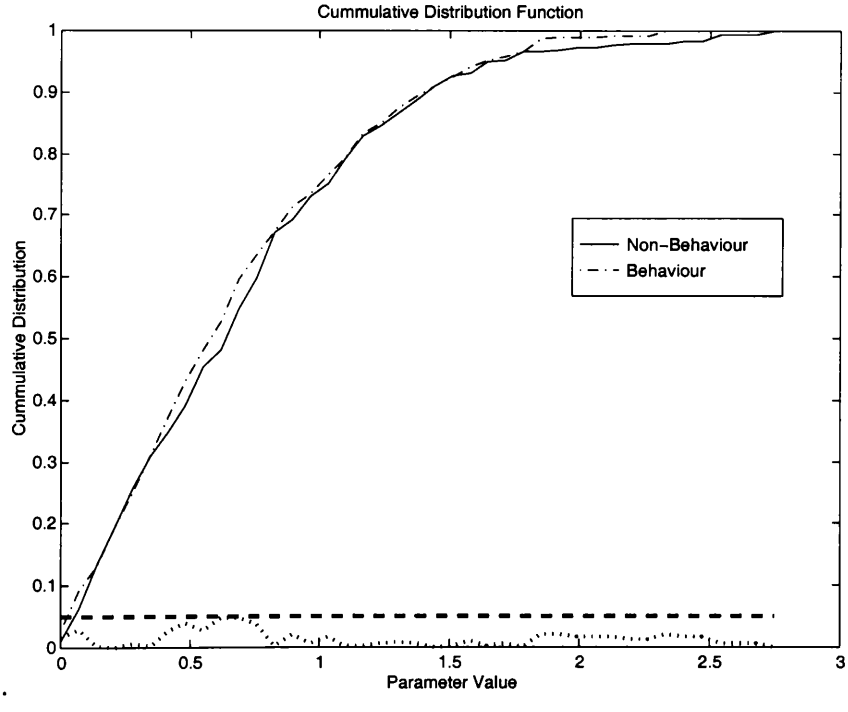

Figure 6.14 Significantly similar B and $\overline{B}$ model responses observed for parameter Set 2, parameter K_1.

These results indicate that individually the parameters are all required to generate a behavioural model response. However, because optimisation has revealed parameter correlation exists, it is probable that the values of parameters individually is unimportant, and that the *combination* of individual parameters is paramount.

The probability of a parameter value producing a B or $\overline{B}$ model response can be obtained by differencing each cumulative distribution. The two histograms that are generated can be subsequently contrasted and compared with the estimated optimal parameters. Understandably, the most significant parameters (highest K-S statistic) have the most contrasting probabilistic histograms as illustrated in Figure 6.15. Where a KS statistic is low, (e.g. μ_5: Set 2) the histograms are less distinguishable. These relationships become more defined as the number of model simulations increases.

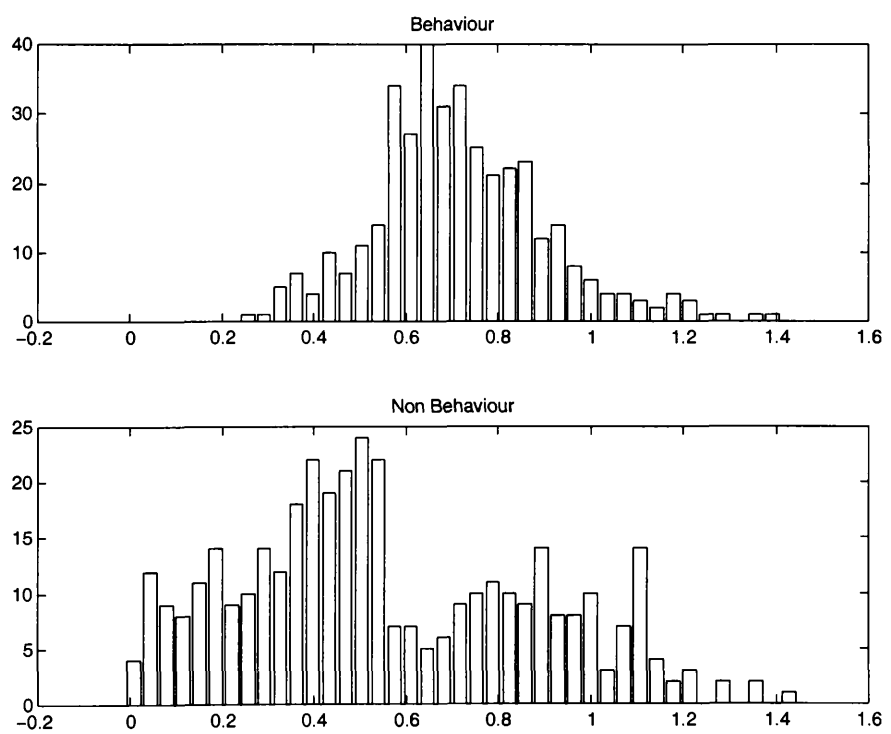

Figure 6.15 Dissimilarity between B and $\overline{B}$ probabilistic histograms, for parameter γ_2 (Set 2).

Except for α and μ_1, the histograms and cdf's of correlated parameters generally have very wide distributions. Generally, the optimal parameter values from Set 2 all have the highest frequency of estimation. However, this was not the case for parameter Set 1. In addition, the results often indicate that the optimal estimated parameter value has an equal probability of generating a B and $\overline{B}$ model response.

From these results, we can see that GSA has been unable to provide a definitive message in this example. It seems that, in order to obtain a greater understanding of the nonlinear parameter interactions that occur within complex models, such as the TCE model, further research is required in the domain of multivariate statistics.

It must be concluded, therefore, that GSA, based on the estimated covariance matrix of the model parameters, has not been able to identify any individual sections or parameters of the model that are particularly responsible for the observed uncertainty in the model output responses. Consequently, the results cannot help to further refine the optimisation of the model. It has also highlighted the fact that simpler deterministic approaches to sensitivity analysis will probably give misleading and inappropriate indications of parameter significance, since they do not take into account the parameter collinearity.

6.7 PARAMETER ESTIMATION BY MAXIMUM LIKELIHOOD

Further TCE model parameter estimation was carried out using the Maximum Likelihood (ML) method outlined in Chapter 2. ML is based on one-step ahead predictions, calculated for nonlinear models using the Linearised Kalman Filter (LKF). The LKF, in its prediction-correction form, recursively estimates the state vector of the system and corrects this estimate as new data becomes available. In other words, it can be seen as a basic data assimilation tool.

The TCE model parameters were estimated by ML using the CTLSM software package (see Chapter 2). The CTLSM software required the following components to be specified:

1. The system and measurement equations $\boldsymbol{f}$ and $\boldsymbol{h}$ (e.g. equations 6.12-6.13);
2. the partial derivatives (Jacobians) of the state and measurement equations;
3. the inputs into the model;
4. the states to be estimated;
5. the parameters to be estimated and their maximum and minimum values.

Calculating the required partial derivatives of $\boldsymbol{f}$, with respect to the parameters, involved an extremely complex procedure, even for the relatively simple TCE model differential equations. If parameters of a more complex model were to be estimated (e.g. Fasham, 1990), calculation of the partial derivatives could become a distinct problem.

Experience using the CTLSM software has showed that the initiating conditions for the model are crucial for the estimation routine to converge. The quasi-Newton algorithm used to generate the preliminary parameter estimates is gradient based and, therefore, it is susceptible to local rather than global convergence.

In light of optimisation difficulties encountered in the Section 6.5, *simulated* rather than observed data were used to evaluate the ML method of optimisation. Nitrate, phytoplankton, zooplankton and NPP output series were generated by simulating the TCE model with parameter values defined in Table 6.6. The ML optimisation was initiated using parameter values close to those used to generate the simulated series. To facilitate a valid evaluation of the ML optimisation procedure, the model was also optimised using simulated data by the nonlinear least squares method.

6.7.1 ML estimation results

The CTLSM estimation routine converged, finally, but only after 36 hours of CPU time on a Sun SPARC ULTRA workstation. This excessive convergence time seems to be a fundamental limitation of this estimation procedure and, in its present configuration, is highly unsuitable for on-line data assimilation purposes. This time would also increase significantly if the parameters of the model were to be estimated as time variable, which is a likely consideration in practical applications. The final ML parameter estimates are presented in Table 6.6 and as shown in Figure 6.16, the estimated model output series are identical to the simulated model output.

Table 6.6 Parameters estimated using the ML method.

Parameter	Simulated	ML Estimated	Std Error
K_1	0.1385	0.0391	0.0154
μ_1	0.006	0.00263	1.935e-4
Po	0.0827	0.0828	4.2368e-5
K_3	0.4388	0.429	4.0542-3
g	1.9962	2.0113	0.01677
γ_2	0.5809	0.5644	1.1834e-3
μ_5	0.0630	0.06277	1.0262e-4
m	2.1875	2.1112	3.5892e-3
γ_5	0.5859	0.5320	0.02582
γ_4	0.3210	0.2379	0.0186
α	0.05	0.05107	2.144e-4

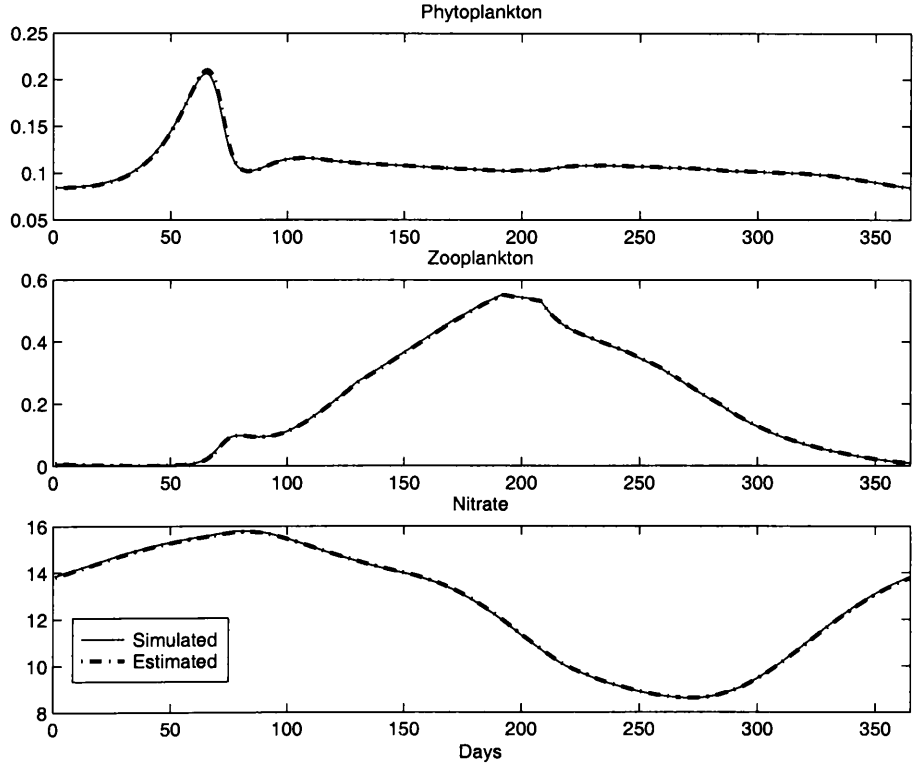

Figure 6.16 Simulated and ML estimated model output

The ML estimated parameter values are generally very similar to those used to generate the simulated model output, (columns two and three in Table 6.6), although there are marked exceptions: K_1, μ_1, γ_5, and γ_4. Table 6.7 presents the cross correlation matrix of the ML estimated parameters and indicates that significant correlation exists between the parameter pairs, K_1:μ_1 and γ_5:γ_4. This correlation accounts for why the estimated values of these parameters deviate from the values used to simulate the model, although an identical model output is maintained.

Table 6.7 Cross correlation matrix of ML estimated parameters

	K1	μ1	Po	K3	g	γ2	μ5	m	γ5	γ4	α
K1	1	0.7	0.44	-0.09	-0.02	-0.33	-0.46	-0.23	0.01	0.02	0.82
μ1	0.7	1	0.38	-0.19	-0.08	-0.49	-0.37	-0.3	-0.03	-0.03	0.89
Po	0.44	0.38	1	-0.48	-0.3	-0.7	-0.82	0.02	-0.03	-0.06	0.44
K3	-0.09	-0.19	-0.48	1	0.98	0.07	0.08	0.09	0.03	0.02	-0.1
g	-0.02	-0.08	-0.3	0.98	1	-0.14	-0.1	0.09	0.03	0.01	0.04
γ2	-0.33	-0.49	-0.7	0.07	-0.14	1	0.88	-0.04	0.01	0.07	-0.6
μ5	-0.46	-0.37	-0.82	0.08	-0.1	0.88	1	-0.07	0.01	0.06	-0.49
m	-0.23	-0.3	0.02	0.09	0.09	-0.04	-0.07	1	0.12	0.03	-0.1
γ5	0.01	-0.03	-0.03	0.03	0.03	0.01	0.01	0.12	1	0.98	0.01
γ4	0.02	-0.03	-0.06	0.02	0.01	0.07	0.06	0.03	0.98	1	-0.01
α	0.82	0.89	0.44	-0.1	0.04	-0.6	-0.49	-0.1	0.01	-0.01	1

In contrast, nonlinear least squares was unable to optimise the model. Although, initiating the optimisation procedure in an identical manner to the ML approach, when the algorithm converged, the values of the model parameters were substantially different from those used in the simulation and the model output series were far from identical to the simulated series.

Whilst these results highlight the superior parameter estimation qualities of ML, additional analysis revealed further limitations of the CTLSM ML procedure. The poor robustness of this parameter estimation approach becomes clear when the routine

is initialised with the parameter values that are not close to the known optima and when the model is optimised based on *measured*, rather than simulated data. The ML procedure (in the CTLSM format) failed badly in both situations. Due to the nature of the gradient based (see Chapter 2) quasi-Newton algorithm, used to generate parameter estimates within the CTLSM software, if inappropriate initial parameter values are specified (i.e. causing the model to go unstable) the estimation routine cannot easily track towards a minima. An algorithm that generates parameters incorporating a random component, (i.e. genetic algorithm) would help alleviate this problem. Modifications of the CTLSM or alternative ML procedures to include such methods could form a suitable topic for future research.

The CTLSM software would not converge when applied to the *measured* data and crashed completely midway (after 18 hours) through the estimation procedure. Repeated attempts to achieve convergence by defining different initial conditions failed and, in the end, had to be abandoned due to time constraints. An alternative, computationally effective method would be to use the continuous-discrete Kalman filter mechanisations. The differential equations would form the system equation (as with the LKF) and integration would yield the one step ahead predictions. The use of the continuous-discrete Kalman filter would also mean that irregular sampled data could be utilised, so enhancing the practical utility of this approach to ML estimation and data assimilation. The results from this ML estimation analysis indicate that such modifications would be essential to optimise models with a larger and more complex structure.

6.8 MODEL LINEARISATION AND REDUCTION

With the continued evolution of modern super computers, there is a tendency for more detailed processes to be included in deterministic models, increasing still further their complexity and associated uncertainty. Some researchers, however, are recognising that model complexity is often unnecessary and detrimental to the aims of the project. Complex environmental models can be effectively simplified, using various techniques, whereby the dominant behavioural modes of the dynamic system can be identified and represented by a series of simple reduced order linear models. Most topically, the complexity of global climate models have been successfully addressed in recent research by Young *et al.*, (1996), Hasselmann *et al*, (1997), Hasselmann (1997) and Joos *et al,* (1996). Such simplification of complex nonlinear oceanic ecosystem models, retaining the dominant modes of the system, would be particularly advantageous for use in data assimilation frameworks, particularly in view of the problems identified in Section 6.7.1.

Simplification of nonlinear models can be achieved by model linearisation and reduction using a number of techniques. These include traditional Taylor series linearisation and state space model reduction, as well as data-based identification/estimation methods that linearise and reduce the model in one step. This latter data-based approach is particularly advantageous for use with complex high order models, where traditional linearisation approaches tend to fail. Both methodologies are based upon Lyapunov's proof, that there is a region around an equilibrium point, however small, that can be described by a linear model. There is, of course, the possibility that the region of linear behaviour may be very small indeed and

so the large number of linear models are required to characterise the whole operating range of the model, would make linearisation an unsuitable option. However, practical experience suggests that this is not normally encountered and a 'piece-wise linear' model quite often provides a good approximation to many nonlinear models. Previous experience (Young *et al.*, 1996; Lees, 1996) has found that linear models can often successfully characterise the nonlinear dynamics over a wide operating range of the model. Where a constant parameter linear model may fail to approximate the nonlinear model behaviour, it is probable that a time varying parameter or piece-wise linear model can be successfully utilised. Furthermore, any system nonlinearity revealed in the parameter estimates may exhibit state dependence which can be subsequently identified following DBM techniques as demonstrated in Chapters 3, 4 and 5.

6.8.1 Classical linearisation and model reduction

Nonlinear models characterised by differential equations can be linearised by the technique of calculating the first term of the Taylor series expansion for a small deviation in the inputs and states around a steady state operating point. A linear model is therefore formed by calculating the rate of change of the state derivatives (Jacobians) and the rate of change of the outputs with respect to the inputs at a specified steady state operating point. The resulting continuous-time model can be written in state space format:

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} \tag{6.14}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \tag{6.15}$$

For any reasonably defined nonlinear model, equations (6.14-6.15) can be calculated using the LINMOD function available in MATLAB/SIMULINK, where the nonlinear model is written either as an s-function or as a SIMULINK block diagram (MathsWorks, 1992). Subsequent to linearisation, it is sometimes desirable to reduce the order of the model in order to simplify the model function. Such model reduction can be achieved by removing those states from the model which are perceived as not significant in their effect on the dominant mode of model behaviour and solving the state space equations when these derivatives are set to zero.

6.8.2 Data based approach to model order reduction and linearisation

The data based approach to model linearisation and reduction developed by Young *et al.*, (1996) consists of two stages. Firstly, the steady state condition or equilibrium point for constant inputs of the simulation model are defined. This can be achieved either by simulating the model until steady state conditions are reached, or using the MATLAB/SIMULINK constrained optimisation function TRIM. Secondly, each model input variable is then individually perturbed around its constant value using a signal that will sufficiently or persistently excite the system, i.e. a Pseudo-Random Binary Signal (PRBS) or some alternative, depending on the example. Based on this input-output data series, a reduced order multi-variate linear transfer function model is identified and estimated using the advanced statistical SRIV routine.

The complete reduced order model obtained in this manner takes the form of a m input and p output multi-input multi-output (MIMO) model where each individual input-output element takes the following single-input single-output (SISO) form:

$$y_j(k) = \frac{B_{ij}(z^{-1})}{A_{ij}(z^{-1})} u_i(k) \tag{6.16}$$

Here, $u_i(k)$ and $y_j(k)$ are the ith input and jth output variables, respectively, characterised by the polynomials $A_{ij}(z^{-1})$ and $B_{ij}(z^{-1})$ of order *nij* and *mij* in the backward shift operator. If a pure time delay, δ, effects the system, this is accommodated by setting the δ leading coefficients of $B_{ij}(z^{-1})$ to zero. The model is identified and parameters estimated using the SRIV algorithm. The resulting parsimonious model should capture the dominant dynamics of the system.

6.8.3 Results of classical and DBM model linearisation

The TCE model has five exogenous inputs each effecting the model response to different extents. The magnitude of each input changes quite considerably over its annual cycle, effectively equating to an infinite number of model operating points. The mixed layer depth (M) has the most significant effect on the model output and to evaluate whether linearisation of the TCE model is possible at this preliminary stage, the model was linearised with respect to M only. All other input variables were set constant at their respective mean annual values as shown in Table 6.8.

Table 6.8 Values of input variables for model linearisation and order reduction

Input Variable	Constant Input
Total Irradience (I)	55.70 Wm^{-2}
Mixed Layer Temperature (mt)	8.14° C
Day Length (2τ)	0.25
Nitrate Conc. Below the Mixed Layer (N_o)	13.00

Table 6.9 Steady state conditions (SSC) calculated by continuous simulation of the TCE model and by constrained optimisation, MATLAB/SIMULINK TRIM function.

Constant Input M	SSC calculated by continuous simulation of the TCE Model and by constrained optimisation (TRIM function).		
	N	P	Z
20	12.639	0.2035	0.963
30	12.571	0.2035	0.921
40	12.530	0.2035	0.814
50	12.508	0.2035	0.707
60	12.495	0.2035	0.616
70	12.489	0.2035	0.538
80	12.486	0.2035	0.474
90	12.486	0.2035	0.422
100	12.487	0.2035	0.379

It is necessary to linearise the TCE model over the annual variation of input M, which ranges from ~18-100m. Both linearisation methods require the steady state conditions (SSC) to be determined at different operating points of the model (i.e. over the range of input M). For comparative purposes, the SSC of the TCE simulation model were calculated using the two contrasting methods outlined in Section 6.8.3. As expected, both methods gave exactly the same value for the steady state conditions at each constant input (M) operating point, as shown in Table 6.9.

The results from utilising the LINMOD MATLAB function to linearise the TCE model were disappointing. At each operating point and for each model state, the LINMOD function determined a linear state space model. However, examination of each of these state space models highlighted (unexpectedly) the fact that each model failed to replicate the behaviour of the nonlinear TCE model. Thorough investigation could not determine why LINMOD was unable to linearise the TCE model and consequently, this method of linearisation was abandoned.

The data-based approach to linearisation was carried out using a mixed-layer depth input (M) covering the *complete* 18-100m operating range. To this effect, the TCE model was simulated using the measured mixed layer depth series to generate P, Z and N outputs. The SRIV algorithm was subsequently utilised to identify and estimate 3 *linear* TF models based upon this single input (M) and 3 output data series. The results of the model identification/estimation are presented in Table 6.10. The model identification criteria clearly demonstrate that the individually estimated linear models are parsimonious and provide a very good fit to the data over the complete operating range, which is additionally shown visually in Figure 6.17. Whilst the exact dynamics of the nonlinear simulation model are not replicated by the linear model, for data assimilation purposes (e.g. the Kalman filter), the benefits of the simplified model structure far outweigh the resulting slight compromise in model fit. Furthermore, it was discovered that by decreasing the magnitude of the mixed layer depth M input into the nonlinear simulation model by 30%, linear models could be identified which gave an almost identical fit to the simulation model outputs. This highlights the true operating range of the linear models.

Table 6.10 Best linear TF models identified between input series M and nonlinear simulation model output series N, Z, and P.

State	Model Order	*YIC*	R_T^2	*AIC*
N	[2,2,4]	-10.710	0.964	-8.686
P	[2,2,0]	-8.337	0.946	-13.645
Z	[2,2,4]	-10.373	0.960	-5.036

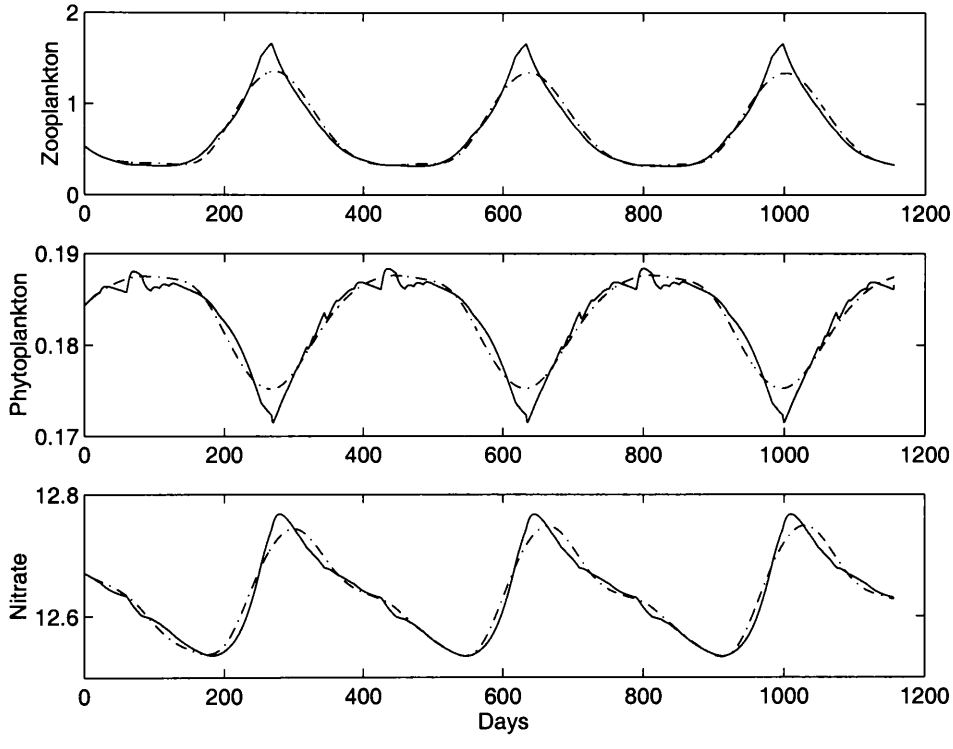

Figure 6.17 Linearised (dotted line) and nonlinear (full line) models of N, P, Z.

Moreover, linear models were identified at each operating point, for P, Z, and N. For each separate output, the best identified TF models showed a consistent structure over the *complete* range of M, which enables the models to be combined, to form a piece-wise linear time varying parameter TF model. Whilst a slightly enhanced model fit was obtained by adopting this approach, the magnitude of this increase was not such to

outweigh the disadvantages of identifying a model with greater structural complexity and associated uncertainty. As a result, the piece-wise linear TF modelling approach was rejected.

The encouraging results from this preliminary analysis indicate that the nonlinear simulation model could be fully linearised with respect to all the model inputs forming a MIMO TF model of the form introduced in equation (6.16). Additional analysis also suggests that two of the inputs (mt and τ) have such a minimal effect on the model output that they could be eliminated from the fully linearised MIMO TF model. Although these changes may not be significant for this relatively simple ecosystem model, for larger higher-order models, it should prove very worthwhile going through this procedure. For example, a high dimensional model reduction achieved in this manner by Young *et al.*, (1996) of the Enting and Lassey global carbon cycle model suggests that a high-order oceanic model could be reduced considerably by combined linearisation and reduction by data-based methods.

6.9 CONCLUSION

This chapter has presented the results of a preliminary study investigating the nature of a deterministic nonlinear oceanic ecosystem model. Oceanic ecosystem models have been developed as tools to help understand plankton population dynamics and to assess how the ecosystem may react to global climate change. The models can often be complex and over-parameterised, which can lead to the model outputs and the resulting interpretation of these outputs, to be highly uncertain. As a prelude to a detailed assessment of a complex higher-order ecosystem model, a three component

nitrate-phytoplankton-zooplankton ecosystem model (Matear, 1995) has been evaluated.

Model parameter optimisation was carried out using nonlinear least squares and a genetic algorithm. The outputs generated from the optimised model provide a much better fit to the data than has been previously obtained (c.f. Matear, 1995). The optimised parameter covariance matrix shows evidence of significant model parameter cross-correlation or collinearity, indicating that parameters could possibly be eliminated from the model. Generalised Sensitivity Analysis was used to determine this eventuality, identifying whether any of the model parameters have an insignificant affect on the model responses. The results indicated that *all* the TCE model parameters were required to generate behavioural responses. However, the analysis highlighted that the individual parameter values were not significant, rather, the resulting value of the parameters in combination.

Further optimisation of the model was carried out using the stochastic Maximum Likelihood (ML) approach. Parameters are estimated using ML based upon one step ahead predictions calculated, in this instance, using the Linearised Kalman filter (LKF). If the model parameters were updated in this manner online, this procedure could be classed as *data assimilation.* ML estimation of the TCE model was undertaken using the CTLSM software which incorporates the discrete-time LKF. Whilst CTLSM provided good ML estimates of the model parameters (superior to those obtained using nonlinear least squares) based upon *simulated* data it proved to be unable to optimise the model using the observed data series. Furthermore, the computation times involved in the estimation procedure prohibit it from being used as

a data assimilation tool. Consequentially, if the advantages of model parameter estimation by ML are required in the future, the estimation routine would have to be rewritten (probably using MATLAB) using the continuous-discrete LKF. In this alternative form, the computational time needed for estimation would be significantly decreased and it would additionally avoid the necessity for incoming data measurements to be equally spaced, which has immediate advantages for data assimilation purposes.

The nonlinear TCE model was effectively linearised using the DBM approach to model linearisation with respect to the mixed layer depth, the most significant input into the model. For nitrate, phytoplankton and zooplankton, 2nd order linear TF models were identified which characterised the dominant nonlinear dynamics of the TCE model. These initial results suggest that the nonlinear model could be fully linearised to form a multiple-input-multiple-output TF model. Whilst the linear model may not have significant advantages over the TCE model, linearisation of a more complex higher order ecosystem model would have distinct benefits for data assimilation.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATAIONS FOR FUTURE RESEARCH

The chapter is structured into two sections: the first section draws together the main contributions of the thesis; and the second section discusses future research possibilities involving the development and use of the DBM methodology in environmental systems.

7.1 DATA BASED MECHANISTIC MODELLING

Throughout this thesis, the DBM modelling approach has provided the underlying philosophy for modelling nonlinear environmental systems. Despite the difficulties of modelling complex nonlinear systems, the DBM approach, which exploits the advantages of transfer function models, has proven to be an excellent means of characterising the dominant behavioural modes of the systems, from only limited time series data. Further, the nonlinear aspects of the observed system behaviour have been identified in an objective manner, ensuring that the modeller's pre-conceptions of the system nonlinearity are not over-emphasised in the identification of the models

structure. This relative objectivity in modelling terms has been made possible by the collective application of the Fixed Interval Smoothing (FIS), Simplified Refined Instrumental Variable (SRIV) algorithms, and the use of the Maximum Likelihood (ML) method in final model estimation.

7.2 APPLICATIONS OF DBM MODELLING

Whilst the results of Nicholson's famous sheep blowfly experiments (1950; 1954; 1957) have been extensively studied by many scientists, the complete egg-blowfly population dynamics have never been satisfactorily modelled. This has provided the impetus for the application of the objective DBM modelling methodology to the analysis of the complete egg-blowfly system.

The nature of the egg-blowfly feedback nonlinearity is particularly crucial in the development of a successful model and its form has been investigated using statistically rigorous methods applied directly to the time series data with a minimum of prior assumptions. Time varying parameter estimation has yielded a characteristic nonlinear relationship, which exhibits clear state dependence. The resulting fully parameterised 'state dependent parameter' model provides a very good explanation of the egg and adult blowfly data. Unknown to the author until after the modelling study was completed, previous researchers (Gurney *et al.*, 1980; 1983) have deduced the same form of the feedback nonlinearity through a heuristic approach, by visual inference from the data but without any statistical identification and verification. However, in most modelling applications, this 'hypothetico-deductive' approach would not be as successful, since the average modeller's insight about the system is

unlikely to be as effective as that of Gurney's. As such, this ecological example clearly demonstrates the true potential and efficacy of the DBM methodology for modelling nonlinear systems.

It is well known within the field of hydrology that the rainfall-flow relationship is inherently nonlinear and fundamentally affected by catchment antecedent soil moisture conditions. Effective rainfall-flow models must address this nonlinearity but, since time series data quantifying soil moisture conditions are rarely available, modellers have generally tackled this problem by generating a 'soil moisture surrogate' in either a conceptual or deterministic manner. In the case of the Lancaster DBM rainfall-flow model, the form of the soil moisture nonlinearity is objectively identified and parameterised from the rainfall and flow time series directly, using time varying parameter and state dependent parameter estimation techniques. The availability of time series measurements of the soil moisture status from a Swiss mountain catchment has enabled the effectiveness of the Lancaster soil moisture surrogate to be determined for the first time. From two independent data series, the soil moisture surrogate generated by the Lancaster DBM model has been shown to capture the dominant dynamics of the measured soil moisture variables and exhibit greater correlation than an alternative conceptualised soil moisture surrogate generated by the IHACRES model (see e.g. Jakeman *et al.*, 1990a).

These results highlight the efficacy of the Lancaster DBM rainfall-flow model and indicate that the model can be applied with confidence to catchments where extensive soil moisture time series are unavailable. The research has also shown that, where additional descriptive soil moisture data are available, it can be incorporated directly

within the rainfall-flow model, not only to improve the explanation of the data, but also to enhance the model's physical interpretation.

Continuing with the hydrological theme, the DBM approach has been used in Chapter 5 as an exploratory tool to investigate historical reservoir sedimentation at Wyresdale Park catchment, Lancashire. The rainfall-sediment load relationship is extremely complex and nonlinear and, as yet, it has not been modelled effectively in a parsimonious manner. Two years of field data were available to investigate the rainfall-sediment relationship at Wyresdale and, whilst the identified system nonlinearity was not fully characterised in state dependent terms, the resulting identified nonlinear model provides a sufficiently good explanation of the time series data for hindcasting purposes. Coupled with a nonlinear TF model for sediment transmission across Wyresdale reservoir, the total DBM model has been shown to provide an effective means of reconstructing historical sediment yields and key sediment profiles within the lake deposits.

In this application, it should be stressed that the full potential of the analysis has also been restricted by the frequency and length of time series data available for model calibration and for hindcasting. Despite the limitations of this particular study, Chapter 5 has introduced a methodology that is able to utilise short periods of time series data to determine more information about historical catchment sediment yields than would otherwise be available from the contemporary records alone.

The deterministic oceanic ecosystem simulation model of Matear (1995) has provided a platform for a preliminary study investigating model order reduction, due to over-

parameterisation, with the overall goal of incorporating the model within a data assimilation framework. By considering the model in stochastic form, statistical DBM techniques have been utilised to expose any poorly identifiable parameters of the model. In its present study, the model parameters have been successfully optimised based upon one-step ahead predictions, generated using the linearised Kalman filter. The online utilisation of this approach can lead to the development of data assimilation methods based on linearised but time variable parameter models. The research has demonstrated the initial feasibility of this approach but its further development is a topic for future research.

7.3 FURTHER RESEARCH

This thesis has highlighted many promising areas for future nonlinear DBM research, both in terms of applying the DBM philosophy to other nonlinear systems and also developing the nonlinear DBM methodology further.

Whilst the rainfall-sediment model adequately describes the rainfall-sediment relationship, future research should aim to examine the specific form of the nonlinear transformed rainfall component of the model in more detail. It is hoped that further analysis of the rainfall-sediment nonlinearity using time-varying and state dependent parameter analysis will reveal a more definitive state dependent relationship which will not only increase the model fit to the data but also enhance the model's physical interpretation. Additional data from the Wyresdale and other catchments would assist in this aim. The time varying/state dependent parameter approach to identifying system nonlinearities could also be utilised readily in other hydrological systems, e.g.

in the identification and estimation of the nonlinear relationship between flow and stage, which would have the advantage of providing standard errors on the rating curve. Additionally, this methodology could be used in the field of hydrometry; for example, in instrumentation calibration.

It has been concluded that the DBM approach to modelling historical sediment loads shows great potential, although it would be more ideally suited to catchments where the dynamics are more distinct in daily sampled data. For this reason, it is hoped that, in the future this methodology will be applied to much larger catchments. Further, larger catchments may make the validation of a hindcasted synthetic sediment sequence simpler and more reliable. Sediments in reservoirs draining a larger area may contain more distinctive sediment horizons, which may be more easily identifiable in core analysis, especially when using new methods for core sequencing such as x-ray densiometry (Cooper and O'Sullivan, 1998).

The encouraging results from Chapter 6 highlight, therefore, that even in a simple oceanic ecosystem model, there is some evidence for model over parameterisation. Further, it has been shown that the nonlinear model can be efficiently linearised using the DBM approach. The opportunity arises therefore, to investigate parameter identifiability, combined statistical linearisation, and model order reduction in more complex, higher order, oceanic ecosystem models where it is likely that model over-parameterisation is more significant. It is expected that the benefits of model order reduction for forecasting and data assimilation will be far more consequential for these larger more complex models.

In Chapter 6 the stochastic ML method of parameter estimation has been utilised to optimise the oceanic ecosystem model based upon *all* the available time series data, i.e. in an off-line manner. Future work should aim to implement an on-line version of the ecosystem model and, therefore, achieve *data assimilation* in its true form. To facilitate this implementation, a new ML estimation routine would be need to be developed, using MATLAB, which incorporates state estimation based on the continuous-discrete linearised Kalman filter. As such, the ecosystem model can remain defined in its continuous form but it will have the advantage of being updated based upon discretely measured data that does not have to be sampled equally in time.

Further impetus for general algorithm development has come from research in Chapter 3. Here, a novel method of estimating states in nonlinear models based on state dependent parameter estimation has been introduced, again based on ML estimation. The identified state dependent time varying parameters, which characterise the system nonlinearity, can be solved at each recursion prior to state prediction and correction being undertaken. In this manner, the *linear* form of Kalman filter can be utilised rather than the computationally more intensive *linearised* Kalman filter, providing a more efficient means of optimisation. Future research should aim to develop a generic algorithm that would allow any state dependent parameter definition to be incorporated within the Kalman filter as part of a ML based nonlinear model parameter estimation routine.

The results presented in this thesis have shown how the DBM modelling methodology can be used to effectively represent nonlinear environmental systems. The DBM approach and associated mathematical algorithms are truly generic and are capable of

providing simple models that are physically interpretable. It is hoped that, in the future, the full potential and benefits of the DBM approach to modelling are realised by other researchers and are applied to a wide range of nonlinear systems.

APPENDIX 1

AUTHORS PUBLICATIONS

Fawcett, C. P, Young, P. C. and Feyen, H, (1997). 'Validation of data-based mechanistic non-linear rainfall-flow model, *Proceedings of the ICSE-97 Twelfth International Conference on Systems Engineering*, **1**, 252-258.

Fawcett, C. P, Young, P. C. and Feyen, H, (1997). 'A data-based mechanistic non-linear rainfall-flow model: Validation using soil moisture data, *19th British Geomorphological Research Group (BGRG) Postgraduate Symposium*,

Price, L.E., Fawcett, C.P., Rowan, J.S. and Young, (1999). 'Modelling reservoir sedimentation using a data-based mechanistic (DBM) modelling approach: Wyresdale Park Reservoir, Lancashire, UK'. *Hydrological Sciences Journal. In press.*

Rowan, J. S. Price, L. E., Fawcett, C. P. and Young, P. C. (1999) 'Reconstructing sedimentation trends in reservoirs using a data-based mechanistic modelling approach'. *Submitted to Physics and Chemistry of the Earth.*

REFERENCES

Beck, M. B. (1987). 'Water quality modelling: a review of the analysis of uncertainty', *Water Resources Research,* **23**(8), 1393-1442.

Beck, M. B. (1993). 'Uncertainty, identifiability, and predictability in environmental models'. In Young, P. C (Ed.), *Concise Encyclopaedia of Environmental Systems*, Pergamon, London, 638-644.

Beck, M. B. and Young, P.C. (1976). 'Systematic identification of DO-BOD model structure. *Journal of Env. Eng. Div., ASCE*, **102**, 909-927.

Beck, M. B. and van Straten G. (1983). *'Uncertainty and forecasting of water quality'*, Springer, Berlin.

Beven, K. J. (1987). 'Towards a new paradigm in hydrology' *International Association of Hydrological Sciences Publication*, **164**, 393-403.

Beven, K. J. (1989). 'Changing ideas in hydrology-The case of physically based model,' *Journal of Hydrology*, 157-172.

Billings, S. A. and Voon, W. S. F. (1986). 'Correlation based model validity tests for nonlinear models', *Int. J. Control*, **44**, 235-244.

Box, G. E. P and Jenkins, G. M. (1970). *'Time Series Analysis Forecasting and Control'*, Holden-Day San Fransisco, California.

Box, G. E. P., and Jenkins, G. M. (1976). *'Time Series Analysis Forecasting and Control'*, 2nd Edition, Holden Day, San Fransisco, California.

Brasseur, P. (Ed.) (1995). 'Data assimilation in marine science', *Journal of Marine Systems*, **6**(1-2), 175 pp.

Brillinger, D. R., Guckenheimer, J., Guttorp, P. and Oster, G. (1980). 'Empirical modelling of population time series data: The case of age and density dependent vital rates', *Lectures on Mathematics in the Life Sciences*, **13**, 65-90.

Chen, J. and Beck, M. B. (1993). 'Modelling, control and online estimation of activated-sludge bulking', *Water Science and Technology*, **28**(11-12), 249-256.

Chen, T. H, Hornberger, G. M., Jakeman, A. J. and Swank, W. T. (1995). 'The performance of different loss models in the simulation of streamflow, *Environmetrics'*, **4**, 479-484.

Curr, R. H. F. (1985). 'Magnitude and frequency of fluvial sediment transport determined from recent lake sediment cores', Chapter 6 in: I.D.L. Foster, A. M. Gurnell, and B. W. Webb (Eds), *Sediment and Water Quality in River Catchments,* Wiley, Chichester.

Duck, R. W. (1996). 'Application of the QDa-md method of environmental discrimination to particle size analyses of fine sediments by pipette and SediGraph methods – Reply', *Earth Surface Processes and Landforms*, **21**(5), 479-481.

Eigbe, U., Beck, M. B., Wheater, H. S., Hirano, F. (1998). 'Kalman filtering in groundwater flow modelling: problems and perspectives', Stochastic Hydrology and Hydraulics, **12**(1), 15-32.

Evans, G. T., and J. S. Parslow. (1985). 'A model of annual plankton cycles, *Biological Oceanography'*, **3**(3), 327-347.

Fasham, M. J. R. (1985). 'Flow analysis of materials in the marine euphotic zone'. In Ulanowicz, R. E. and Platt, T. (Eds), *Ecosystem Theory for Biological Oceanography, Can. Bull. Fish. Aquat. Sci.*, **213**,139-162.

Fasham, M. J. R. (1993). 'Modelling the marine biota'. In: M. Heimann (Ed), *The Global Carbon Cycle,* Springer, Berlin, pp. 457-504.

Fasham, M. J. R. and Evans, G. T. (1995). 'The use of optimisation techniques to model marine ecosystem dynamics at the JGOFS station at 47° N, 20° W', *Phil. Trans. R. Soc. London.* **B348** 203-209.

Fasham, M. J. R., Ducklow, H. W. and McKelvie S. M. (1990). 'A nitrogen based model of plankton dynamics in the oceanic mixed layer', *Journal of Marine Research* **48** 591-639.

Fawcett, C. P, Young, P. C. and Feyen, H, (1997). 'Validation of data-based mechanistic non-linear rainfall-flow model, *Proceedings of the ICSE-97 Twelfth International Conference on Systems Engineering*, **1**, 252-258.

Fedra, K. van Straten, G., Beck, M. B. (1981). 'Uncertainty and arbitrariness in ecosystem modelling: A lake modelling example', *Ecological Modelling*, **13**, 87-110.

Feyen, H., Leuenberger, J., Papritz, A., Gysi, S. M., Flehler, H. and Schleppi, P. (1996). 'Runoff processes in catchments with a small scale topography', *Phys. Chem. Earth.*, **21**, 177-181.

Fleming, G. (1975). '*Computer Simulation Techniques in Hydrology*', Elseiver

Fogg, G.E. (1975). 'Primary productivity'. In: Riley, J. P. and J. Skirrow (Eds), *Chemical Oceanography Vol 2*, Academic Press, London, pp 385.

Foster, M. J. (1995). 'Continuous Time Estimation and its Application to Active Mixing Volume (AMV) Models', *Ph.D Thesis, Lancaster University*, Lancaster.

Foster, I. D. L., Dearing, J. A., and Appleby, P. G. (1986). 'Historical trends in catchment sediment yields: a case study in reconstruction from lake-sediment records in Warwickshire, UK'. *Hydrological Sciences Journal*, **31**(3), 427-443

Freeman, T. G. (1981). 'Introduction to the use of CAPTAIN for time-series analysis. CSIRO Division of Computing Research, Computing', *Note No. 43*

Commonwealth Scientific and Industrial Research Organisation, Canberra, Australia.

Frost, B. W. (1991). 'The role of grazing in nutrient rich areas of the open sea', *Limnol. Oceanogr.*, **36(8)**, 1616-1630.

Goldberg, D.E (1989). 'Genetic algorithms in search, optimisation and machine learning', Addison-Wesley, Reading, Mass.

Goodwill, P. (1998). 'Sediment supply and the transmission dynamics in reservoired catchments of the Bowland Fells, Lancashire, UK', (PhD Thesis,) *Lancaster University*, Lancaster.

Goodwill, P., Rowan, J. S. and Greco, M. (1995). 'Sediment routing through reservoirs, Wyresdale Park reservoir, Lancashire, U.K.', *Phys. Chem. Earth*, **20**, 83-190.

Gran, H. H., and Braarud, T. (1935). 'A quantitative study of the phytoplankton in the Bay of Fundy and the Gulf of Maine', *J. Biol. Board Can.*, **1**, 279-467.

Gurney, W. S. C, Blythe, S, P. and Nisbet, R. M. (1980). 'Nicholsons blowflies revisited', *Nature*, London, **287**, 17-21.

Gurney, W. S. C., Nisbet, R. M. and Lawton, J. H. (1983). 'The systematic formulation of tractable single-species population models incorporating age structure', *Journal of Animal Ecology*, **52**, 479-495.

Harmon, R and Challenor, P. (1997). 'A Markov Chain Monte Carlo Method for estimation and assimilation into models', *Ecological Modelling*, **101(1),** 41-59.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, New York.

Hasselmann, K (1997). 'Climate change research after Koyoto', *Nature*, 390, 225-226.

Hasselmann, K., Hasselmann, S., Geiring, R., Ocana, V. and v. Storch, H. (1997) 'Sensitivity study of optimal CO_2 emission paths using a simplified structural integrated assessment model (SIAM)', *Climate Change*, **37**, 345-386.

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

Hornberger, G. M., Beven, K. J., Cosby, B. J., and Sappington, D. E. (1985). 'Shenandoah watershed study: calibration of a topography based, variable contributing area hydrological model to a small forested catchment', *Water Resources Research*, **25**, 2177-2185.

Jakeman, A. J. and Hornberger, G. M. (1993). 'How much complexity is warranted in a rainfall-runoff model', *Water Resources Research*, **29(8),** 2637-2649.

Jakeman, A. J., Littlewood, I. G. and Whitehead P. G. (1990a). 'Computation of the instantaneous unit hydrograph and identifiable component flows with application to small upland catchments', *Journal of Hydrology*, **117**, 275-300.

Jakeman A. J., Ghassemi, F. and Dietrich, C. R. (1990b). 'Calibration and reliability of an aquifer system model using generalised sensitivity analysis', *IAHS Publication*, **195**, 45-51.

Jakeman, A. J., Littlewood, I. G. and Whitehead P. G. (1993). 'An assessment of the dynamic response characteristics of streamflow in the Balquhidder catchment', *Journal of Hydrology*, **145**, 337-355.

Jakeman, A. J., Post, D. A. and Beck, M. B. (1994). 'From data and theory to environmental model: The case of rainfall runoff', *Environmetrics*, **5**,297-314.

Jamart, B. M., Winter, D. F., Banse, K. (1979). 'Sensitivity analysis of a mathematical model of phytoplankton growth and nutrient distribution in the

Pacific Ocean off the north-eastern US coast', *Journal of Plankton Research*, **1**, 267-290.

Jamieson, D. G. and Wilkinson, J. C. (1972). 'River Dee research program', *Water Res. Res.*, **8**(4), 911.

Jarvis, A. J., Young, P. C., Davis, H. J. and Taylor, C. J. (1998). 'An analysis of the dynamic response of stomatal conductance to a reduction in humidity over leaves of Cedrella odorata', *Plant Cell and Environment*, in press.

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

Jernigan, R. R. and Tsokos, I. P. (1980) 'A linear stochastic model for phytoplankton production in a marine ecosystem', *Ecological Modelling*, **10**, 1-12.

Jonach-Clausen, T. (1979). *'SHE, Système Hydrologique Européen, a short description'*, Danish Hydraulic Institute.

Joos, F., Bruno, M., Fink, R. Siengenthaler, U., Stockler, T. F., Le Quere, C. and Sarmiento, J. L. (1996). 'An efficient and accurate representation of complex oceanic and biosphere models of anthroprogenic carbon uptake', *Tellus*, **48B**, 397-417.

Kalman, R. E. (1960). 'A new approach to linear filtering and prediction problems', *Transactions of the ASME-Journal of Basic Engineering*, Ser. D., **82**, 35-45

Kanda, J., Ziemann, D. A., Conquest, L. D. and Bienfang, K. (1990). 'Nitrate and ammonium uptake by phytoplankton populations during the spring bloom in Auke Bay, Alaska', *Estuarine coastal and shelf science*, **30**, 509-524.

Kleissen, F.M., Beck, M. B. and Wheater, H.S. (1990). 'The identification of conceptual hydrochemical models', *Water Resources Research*, **26**(12), 2979-2992.

Knisel, W. G. (Ed.) (1980). 'CREAMS: a field-scale model chemicals, runoff, and erosion from agricultural management systems'.

Kremer, J. N. (1983). 'Ecological implications of parameter uncertainty in stochastic simulation', *Ecological Modelling*, **18**, 187-207.

Labadz, J. C., Burt, T. P. and Potter, A. W. R. 1991. 'Sediment yield and delivery in the blanket peat moorlands of the southern Pennines', *Earth Surf. Proc. Land.*, **16**(3) 255-272.

Landry, M. R., Postel, J. R., Peterson, W. K. and Newman, J. (1985). 'Broad scale distributional patterns of hydrographic variable on the Washington/Oregon shelf'. In: Landry, M. R. and Hickley, B. M. (Eds), *Coastal Oceanography of Washington and Oregon*, Elseiver.

Lees, M. J. (1996). *'Multivariable modelling and proportional-integral-plus control of greenhouse climate'*, Unpublished PhD. Thesis, Lancaster University, UK.

Lees, M., Young, P. C., Ferguson, S., Beven, K. and Burns, J. (1994) 'An adaptive flood warning scheme for the River Nith at Dumfries', *2nd International Conference on River Flood Hydraulics*, 65-77, J. Wiley and Sons, Chichester.

Lemke, A. K. (1990). 'An evaluation of transfer function noise models of suspended sediment concentration', *Prof. Geogr.*, **42**(3), 324-336.

Lemke, A. K. (1991). 'Transfer function models of suspended sediment concentration', *Water Resources Research*, **27**(3), 293-305.

Littlewood, I. G. and Jakeman, A. J. (1994) 'A new method of rainfall runoff modelling and its applications in catchment hydrology'. In Zannetti, P. (Ed), *Environmental Modelling,* Vol. II, pp143-171, Computer Mechanics Publications, Southampton.

Ljung, L. (1987). '*System Identification: Theory for the User.*' Prentice-Hall, Englewood Cliffs, New Jersey.

Loague, K. M. and Freeze, R. A., (1985). 'A comparison of rainfall-runoff modelling techniques on small upland catchments', *Water Resources Research*, **21**, 229-248.

Lynch, J. A., Corbett, E. S., Sopper, W. E. (1979). 'The effects of antecedent soil moisture on stormflow volumes and timing'. *Proceedings of the Third International Symposium in Hydrology, Colorado State University, Fort Collins, Colorado*, 89-99.

Mackereth, F.J.H. (1969). 'A short core sampler for sub-aqueous deposits' *Limnol. Oceanogr.*, **14**, 145-151.

Madsen, H. and Melgaard, H. (1993). 'CTLSM: continuous time linear stochastic modelling', Technical Report, *Institute of Mathematical Statistics and Operations Research, Technical University of Denmark.*

Matear, R. J. (1995). 'Parameter optimisation and analysis of ecosystem models using simulated annealing: A case study at Station P,' *Journal of Marine Research*, **53**, 571-607.

Mathworks (1992). *The Matlab Reference Guide*, The Mathworks Inc. Mass.

May, R. M. (1973). '*Stability and complexity in model ecosystems'*, Princeton University Press, Princeton.

Maynard Smith, J. (1974). *Models in Ecology*, Cambridge University Press, Cambridge.

McKenna, P. (1998) '*Delta Operator: Modelling, Forecasting and Control*, Unpublished PhD Thesis, Lancaster University, UK.

Myrabo, S. (1986). 'Runoff studies in a small catchment', *Nordic Hydrology*, **17**, 335-346.

Nicholson, A. J. (1950). 'Population oscillations cause by competition for food. *Nature*, **195**, 476.

Nicholson, A. J. (1954). 'An outline of the dynamics of animal populations', *Australian Journal of Zoology*, **2**, 9-65.

Nicholson, A. J. (1957). 'The self adjustment of populations to change', *Cold Spring Harbour Symposia on Quantitative Biology*, **22**, 163-172.

Nisbet, R. M. and Gurney, W. S. C. (1982). *'Modelling Fluctuating Populations'*, Wiley, Chichester.

Norton, J. P. (1975). 'Optimal smoothing in the identification of linear time varying systems', *Proc IEE*, **122**, 663-668.

Pace, M. L., Glasser, J. E. and Pomeroy, L. R. (1984). 'A simulation analysis of continental shelf food webs', *Marine Biology*, **82**, 47-63.

Parkinson, S. D. (1995). *'The application of stochastic modelling techniques to global climate change'*, Unpublished PhD. Thesis, Lancaster University, UK.

Parkinson, S. D. and Young, P. C. (1998). 'Uncertainty and sensitivity in global carbon cycling modelling', *Climate Research*, in press.

Parsons, T. R., and Lilli, C. M. (1988). 'Comparative oceanic ecology of the plankton communities of the subarctic Atlantic and Pacific oceans', *Oceanogr. Mar. Biol. Annu. Rev.*, 26, 317-359.

Poole, R. W. (1978). 'The statistical prediction of population fluctuations', *Ann. Rev. Ecol. Syst.*, **9**, 427-448.

Price, L. E. (1999). *Imperfect Mixing in Energy and Mass Transport: A Data-Based Mechanistic Approach*, Unpublished PhD. Thesis, Lancaster University, UK.

Priestley, M. B. (1988). *Nonlinear and Nonstationary Time Series Analysis*, Academic Press, London.

Readshaw, J. L. and Cuff, W. R. (1980) 'A model of Nicholsons blowfly cycles and its relevance to predation theory', *Journal of Animal Ecology*, **49**, 1005-1010.

Reed, D. W. (1984). 'A review of British flood forecasting practise', *Institute of Hydrology, Report No. 90.*

Riley, G. A. (1946). 'Factors controlling phytoplankton populations on Georges Bank', *Journal of Marine Research*, **6**, 54-73.

Robinson, A. R., Lermusiaux, P. F. J. and Quincy Sloan III, N. (1999) 'Data Assimilation', In: *The Sea*, In press.

Rose, K. A. (1993). 'Sensitivity analysis in ecological system models', In: Young, P. C., *Concise Encyclopaedia of Environmental Systems*, Pergamon, London, 534-539.

Rowan, J. S., Barnes, S. J. A., Hetherington, S. L., Lambers, B. and Parsons, F. (1995). 'Geomorphology and pollution: the environmental impacts of lead mining. Leadhills, Scotland', *Journal of Geochemical Exploration*, **52**(1/2), 57-65.

Sarmiento, J. L., Slater, R., Fasham, M. J. R., Ducklow, H. W., Toggweiler J. R. and Evans, G. T. (1993). 'A seasonal three-dimensional ecosystem model of nitrogen cycming in the North Atlantic eupotic zone,' *Global Geochemical Cycles*, **7**, 417-450.

Sarmiento, J. L., Fasham, M. J. R., Slater, R., Toggweiler, J. R., and Ducklow, H. W. (1992). 'The role of biology in the chemistry of CO_2 on the ocean'. In: Farrell, M. (Ed), *Chemistry of the Greenhouse Effect*, Lewis Publishing, New York.

Scavia, D. (1993). 'Lake ecosystem modelling uncertainty'. In: Young, P. C. (Ed.), *Concise Encyclopaedia of Environmental Systems*, Pergamon Press, Oxford, 318-320.

Scavia, D., Powers, W. F., Canale, R. P. and Moody, J. L. (1981). 'Comparison of first order error analysis and Monte Carlo simulation in time dependent lake eutrophication models', *Water Resources Research*, 17(4), 1051-1059.

SECOR (1990). The Joint Global Ocean Flux Study: Science Plan JGOFS Report No. 5.

Selby, M. J. (1993). '*Hillslope Materials and Processes*', Second Edition, Oxford University Press, UK.

Sharma, T. C., Hines, W. G. S. and Dickinson, W. T. (1979). 'Input-output model for runoff-sediment yield processes', *J. Hydrol.*, **40**, 299-322.

Sharma, T. C. and Dickinson, W. T. (1980). System model of daily sediment load', *Water Resources Research*, **16**(3) 501-506.

Shaw E. M. (1988) *Hydrology in Practice*, Van Nostrand Reinhold, London.

Sherman, L. K. (1932). 'Streamflow from rainfall by unit-graph', *Engineering News Record,* **108**, 501-525.

Söderström, T and Stoica, P. (1989). *System Identification*, Prentice Hall, London.

Sorooshian, S and Gupta, H. K. 1995. 'Model calibration', Chapter 3, in Singh, V. P (Ed.) *Computer Models of Watershed Hydrology*, Water Resources Publications, Denver, Colorado.

Spear, R. C. (1970). 'The application of Kolmogorov- Renyai statistic to problems of parameter uncertainty in systems design', *International Journal of Control*, **11(5),** 771-778.

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

Steele, J. H. and Henderson, E. W. (1981). 'A simple plankton model', *American Naturalist*, **117**, 676-691.

Steele, J. H. and Henderson, E. W. (1992). 'The role of predation in plankton models', *Journal of Plankton Research*, **14(1)**, 157-172.

Thornton, K. W. (1993) 'Sensitivity analysis and simulation experimentation'. In Young, P. C. (Ed.), *Concise Encyclopaedia of Environmental Systems*, Pergamon, London, 532-534.

Tomovic, R. (1963) '*Sensitivity analysis of dynamic systems*' McGraw Hill, NY.

van Genuchten, M. T. (1991). 'Progress and opportunities in hydrologic research,' 1987-1990, *U.S. Natl. Rep. Int. Union Geod. Geophys. 1987-1990, Rev. Geophys.*, **25**, 189-192.

van Straten, G., (1980). 'Analysis of model and parameter uncertainty in simple phytoplankton models for Lake Balaton', WP-80-139, International Institute for Applied Systems Analysis, Laxenburg, Austria.

van Straten, G., (1983). 'Maximum Likelihood estimation of parameters and uncertainty in phytoplankton models', In: Beck, M. B. and van Straten, G. (Eds), *Uncertainty and Forecasting of Water Quality Models*, Springer-Verlag, Berlin, 157-171.

Walling, D. E. (1977). 'Assessing the accuracy of suspended sediment rating curves for a small basin', *Water Resources Research*, **13**, 531-538.

Wang, G. T., Singh, V. P., Changling, G. and KeXin, H. (1991) 'Discrete linear models of runoff and sediment discharge from the Loess Plateau of China', *Journal of Hydrology*, **127**, 153-171.

Wheater, H. S, Jakeman, A. J., and Beven, K. J. (1993). 'Progress and directions in rainfall-runoff modelling'. In: Jakeman, A.J., Beck, M. B. and McAleer, M. J. (Eds), *Modelling Change in Environmental Systems*, Wiley, London, 99-130.

Whitehead, P. G. and Young, P. C. (1975). 'A dynamic stochastic model of the Bedford Ouse', In: G. C. Vansteenkiste (Ed.), *Computer Simulation of Water Resource Systems*, North Holland, Amsterdam, 417-438.

Whitehead, P. G. and Young, P. C. (1979). 'Water quality in river systems: Monte Carlo analysis', *Water Resources Research*, **15**, 451-459.

Whitehead, P. G. and Hornberger, G. M. (1984). 'Modelling algal behaviour in the River Thames', *Water Research,* **18(8)**, 945-953.

Whitehead, P. G. and Young, P. C. and Hornberger, G. M. (1979). 'A systems model of stream flow and water quality in the Bedford Ouse River. 1 Stream flow modelling', *Water Research*, **13,** 1155-1169.

Wright, R. F., Roelofs, J. G. M., Bredemaier, M., Blanck, K., Boxman, A. W., Emmet, B. A., Gundersen, P., Hultberg, H., Kjonaas, O. J., Moldan, F., Tietema, A., van Breemen, N., van Dijk, H. F. G. (1995). 'NITREX: responses of coniferous forest ecosystems to experimentally changed deposition of nitrogen', *For. Ecol. Managem.* **71**, 163-169.

Young, P. C. (1970). 'An instrumental variable method for real time identification of a noisy process', *Automatica*, **6**, 271-287.

Young, P. C. (1974). 'Recursive approaches to time-series analysis', *Bull. of Inst. Maths and its Applications*, **10**, 209-224.

Young, P. C. (1976). 'Some observations on instrumental variable methods of time series analysis', *Int. Journal of Control,* **23**, 593-612.

Young, P. C. (1978). 'A general theory of modelling for badly defined dynamic systems'. In: Vansteenkiste, G.C. (Ed), *Modelling, Identification and Control in Environmental Systems*, 103-135, North Holland, Amsterdam

Young, P. C. (1984). *Recursive Estimation and Time Series Analysis*, Springer-Verlag, Berlin.

Young, P. C. (1985) 'The instrumental variable method: a practical approach to identification and system parameter estimation'. In Barker, H. A. and Young, P. C. (Eds.), *Identification and System Parameter Estimation*, Pergamon Press, Oxford, 1-16.

Young, P.C. (1988). 'Recursive extrapolation, interpolation and smoothing of non-stationary and nonlinear time series'. In Identification and System Parameter Estimation, H. F. Chen (Ed) 33-44, Pergamon, Oxford.

Young, P.C. (1989). 'Recursive estimation forecasting and adaptive control'. In C. T. Leondes (Ed.) *Control and Dynamic Systems*, Academic Press, San Diego, 119-166.

Young, P. C. 1992. 'Parallel processes in hydrology and water quality: A unified time-series approach'. *J. IWEM*, **6**, 598-612.

Young, P. C. (1993). 'Time variable and state dependent modelling of non-stationary and non-linear time series'. In: Subba Rao, T. (Ed.), *Developments in Time Series Analysis*, Chapman and Hall, London, 374-413.

Young, P. C. (1998a). 'Data-based mechanistic modelling of engineering systems', *Journal of Vibration and Control*, **4**, 5-28.

Young, P. C. (1998b). 'Data-based mechanistic modelling, generalised sensitivity and dominant mode analysis', *Computer Physics Communications*, **115**, 1-17.

Young, P. C. (1999). 'Data-based mechanistic modelling of environmental, ecological, economic and engineering systems', *Environmental Modelling and Software*, **13**, 102.

Young, P. C. and Jakeman, A. J. (1979). 'Refined instrumental variable methods of recursive time series analysis; Part 1 single input, single output systems', *Int. Journal Control*, **29**, 1-30.

Young, P. C., and Jakeman, A. J. (1980). 'Refined instrumental variable methods of recursive time series analysis; Part III: extensions', *Int. Journal Control*, **31**, 741-764.

Young, P. C., and Beven, K. J. (1991). 'Computation of the instantaneous unit hydrograph and identifiable component flows with application to small upland catchments-Comments', *Journal of Hydrology*, **129**, 389-396.

Young, P. C., and Beven, K. J. (1994), 'Data based mechanistic modelling and the rainfall-flow non-linearity', *Environmetrics*, **5**, 335-363.

Young, P. C. and Pedregal, D. (1997). 'Data based mechanistic modelling in System dynamics in economic and financial models'. In: Heij, C. and Schumacher, H. (Eds) *System dynamics in economic and financial models*, Wiley, Chichester.

Young, P.C. and McKenna P.G, (1999). 'Optimal time varying parameter estimation using an instrumental variable smoothing algorithm'. In prepartation

Young, P. C., Naughton, J. J., Neethling, C. J., Shellswell, S. H. (1973). 'Macro economic modelling: a case study'. In: Eykhoff, P. (Ed), *Identification and System Parameter Estimation*, Elsiever, New York, 145-165.

Young, P.C., Chotai, A. and Tych, W. (1991). 'Identification, estimation and control of continuous-time systems described by delta operator models'. In N. K. Sinha

and G. P. Rao (Eds.), *Identification of continuous-time systems*, Kluwer Academic Publishers, Dordrecht, 363-418.

Young, P. C., Parkinson, S. and Lees, M. (1996). 'Simplicity out of complexity in environmental modelling: Occam's Razor Revisited', *Journal of Applied Statistics*, **23(2&3)**, 165-210.

Young, P. C., Jakeman, A. J. and Post, D. A. (1997). 'Recent advances in the data-based modelling and analysis of hydrological systems', *Water Science Technology*, **36(5)**, 99-116.

Yu, S., Jakeman, A. J. and Pittock, A. B. (1996). 'Modelling rainfall-runoff from large catchment to basin scale: The Goulburn Valley, Victoria,' *Hydrological Processes*, **10**, 863-874.