System Identification Theory Approach to Cohesive Sediment Transport Modelling

by

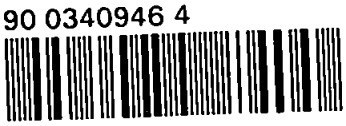
HUIXIN CHEN

Thesis presented for the award of

DOCTOR OF PHILOSOPHY

School of Mathematics and Statistics
Faculty of Technology
University of Plymouth

September, 1997



Dedicated to my mother, my sister, Hong and Le (John)

and to the memory of my father

AUTHOR'S DECLARATION

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

This study was financed with the aid of a studentship from the University of Plymouth.

The work was partially presented at a number of national and international conferences and several papers have been published or will be published in the near future.

Publications:

- H.Chen, A.S.I.Zinober and R.Ruan, 1996.
  Strong Consistency and Convergence Rate of the Parameter Identification for the Bilinear System.
  International Journal of Control, Vol.63 No5. 907-919.

- H.Chen and P.P.G.Dyke, 1995.
  Time Series Models for Concentration of Suspended Particulate Matter and Its Identification.
  submitted to *Estuar. Coastal. Mar Sci.*

- H.Chen and P.P.G.Dyke, 1996.
  Multivariable Time Series Sediment Dynamical Model and Its Identification In Rufiji Delta, Tanzania.
  Applied Mathematical Modelling Vol.20, October, pp.756–770.

- H.Chen and P.P.G.Dyke and H.Zhao, 1997.
  An Minimum Eigenvalue Ratio Test of Product Moment Matrix For Time Series Model Order Estimate.
  submitted *IMA Journal of Control.*

- H.Chen and P.P.G.Dyke, 1996.
  Multivariate Time Series Model For Suspended Sediment Concentration.
  to appear in *Continental Shelf Research*

- H.Chen, J.Blewett, P.P.G.Dyke and D.Huntley, 1997
  Time Series Modelling of Suspended Sediment Concentration on the Holderness Coast.
  submitted to *Journal of Marine Systems*

Presentations:

- H.Chen and P.P.G. Dyke
  Multivariable Time Series Sediment Dynamical Model and Its Identification In Rufiji Delta, Tanzania
  British Applied Mathematics Colloquium, April, 1996, Loughborough.

- H.Chen and P.P.G. Dyke
  Time Series Models For Sediment Transport
  JONSMOD '96, Oslo, Norway

Signed..... Huiyin Chen

Date..... 10/09/97

HUIXIN CHEN

System Identification Theory Approach to Cohesive Sediment Transport Modelling

ABSTRACT

Two aspects of the modelling sediment transport are investigated. One is the univariate time series modelling the current velocity dynamics. The other is the multivariate time series modelling the suspended sediment concentration dynamics.

Cohesive sediment dynamics and numerical sediment transport model are reviewed and investigated. The system identification theory and time series analysis method are developed and applied to set up the time series model for current velocity and suspended sediment dynamics.

In this thesis, the cohesive sediment dynamics is considered as an unknown stochastic system to be identified. The study includes the model structure determination, system order estimation and parameter identification based on the real data collected from relevant estuaries and coastal areas. The strong consistency and convergence rate of recursive least squares parameter identification method for a class of time series model are given and the simulation results show that the time series modelling of sediment dynamics is accurate both in data fitting and prediction in different estuarine and coastal areas.

It is well known that cohesive sediment dynamics is a very complicated process and it contains a lot of physical, chemical, biological and ocean geographical factors which are still not very well understood. The numerical modelling techniques at present are still not good enough for quantitative analysis. The time series modelling is first introduced in this thesis to set up cohesive sediment transport model and the quantitative description and analysis of current velocity and suspended sediment concentration dynamics, which provides a novel tool to investigate cohesive sediment dynamics and to achieve a better understanding of its underlying character.

ACKNOWLEDGEMENTS

I gratefully acknowledge the help of all those who contributed to the completion of this work; in particular the following people.

My deepest thanks to Professor Phil Dyke for giving me the opportunity to register for Ph.D degree and his supervision, guidance, patience and encouragement. I would like to express my sincerest thanks to Dr. Phil James, my second supervisor, for his encouragement and support in many ways especially in fluid mechanics.

I very much appreciate the help of Professor David Huntley whose suggestions helped me to understand the sediment dynamics. I am very grateful to Professor Keith Dyer who provided me with the *in situ* data sets in Rufiji Delta, Tanzanian and his encouragement.

Many thanks to Dr. Julian Stander and Dr. Dave Wright for helpful discussion of some interesting statistical problems.

I would like to thank the School of Mathematics and Statistics, University of Plymouth for providing me such a stimulating research project and pleasant study environment.

I wish to thank my fellow researchers in the School of Mathematics and Statistics and Institute of Marine Science, University of Plymouth particularly Mark, Joanna, Wendy, Yi, An for their help.

Finally I must thank my mother and my sister for their encouragement and moral support, Hong and John who share the busy, colourful and memorable years with me in Plymouth.

Contents

Chapter 1

Introduction

This research has focused upon the mathematical modelling of cohesive sediment transport problem in two areas: one is the current velocity dynamics and the other is suspended particulate matter (SPM) concentration or suspended sediment concentration (SSC) dynamics.

The ability to set up mathematical models of estuarine and coastal sediment dynamics is very important for many research and economic reasons. First, monitoring and control of siltation for the need of navigable harbours and waterway. Second, due to industrialisation and urbanisation, the monitoring and prevention of accumulation of pollutants in close proximity to estuarine and coastal environments becomes a more interested subject. Third, the possibility of global sea level rise has a direct effect on the estuarine and coastal environment. If all sediment transport process can be realistically modelled then predicting the consequences of such an event would be made considerably easier.

Since the most widely used sediment transport flux models are so-called cu-integral (concentration times velocity integral) type of models. Therefore suspended sediment concentration and current velocity are two very important process to model when considering sediment transport.

This thesis is arranged as follows:

The basic concepts and some useful kinds of time series modelling are introduced and

the prediction technique of using time series modelling is shown in Chapter 2.
In Chapter 3, the theoretical background of system identification theory are given. The recursive least squares identification method and its strong consistency and convergence rate are given for different kinds of system or time series model, which is one of the main theoretical contribution of the thesis.
The order determination problem is discussed in Chapter 4 and a new order determination test called minimum eigenvalue ratio test (MERT) is given. The analysis, comparison and simulation results show that it is an improved method for traditional determinant ratio (DR) test.
The description and characteristic properties of cohesive sediment dynamics and its hydraulic numerical modelling are reviewed in Chapter 5.
Chapter 6. gives one layer time series model for current velocity and SPM concentration dynamics in the Holderness Coast, England based on the system identification theory.
In the Chapter 7., the multi-layer time series model for current velocity profile and SPM concentration profile in the Rufiji Delta, Tanzania is described based on the system identification theory.
Finally, conclusions, discussions and suggestions for the further work complete the investigation are presented in the last chapter.

Chapter 2

Time Series Models

2.1 Introduction

A time series is a collection of observations made sequentially in time. (Chatfield 1980, Box and Jenkins 1976, Harvey 1981 and Enders 1995). Typically it consists of a set of observations on a variable y taken at equally spaced interval over time. Examples occur in a variety of fields, ranging from economics, engineering to sediment dynamics. Some general references on the subject are listed in the references section.

1. Examples

We begin with some examples of the sort of time series which arise in sediment dynamics.

(a) Time Series of SPM concentration

A major part of land ocean interaction in the coastal zone consists of the transport of suspended particulate matter (SPM). Here SPM means the mass of the fine solid material in the water with a grain diameter less than 20 μm and greater than 0.4 μm. SPM in the North Sea consists of microflocs of mineral particles and organic matter (mainly detritus) in SPM is about 20 % (Schröder 1988). According to (Eisma & Kalf 1987) in January 1980 about 85 % of the mineral mass was smaller than 20 μm peaking between 2

and 5 μm (Sundermann 1994). Its mass balance is determined by the input from rivers, atmosphere and adjacent seas, by advective and diffusive fluxes, and by deposition and resuspension at the sea bottom. Fig 2.1 shows the time series of SPM concentration in the Tamar Estuary during the high water slack.

(b) Time Series of Current Velocity

The most complete technique for estimating suspended sediment (or SPM) transport is to multiply a current velocity profile by a profile of suspended sediment (or SPM) concentration and integrate the result over the water depth. So current velocity is a very important variable in sediment transport and hydrodynamics, also it is comparatively easy to observe and measure. Fig 2.2 shows the time series of current velocity in the Tamar Estuary during the high water slack.

Suppose one observes such a time series, denoted by y_t for the value at the time period t, over the period from $t = 0$ up to $t = n$, where n means now. In time series forecasting we are interested in making statements about what value the series will take at some future time period $n + h$, where h means hence. Thus the h represents the number of time periods into the future the forecast is looking. If $h = 1$, then the one-step forecasts are being made. For example (see Fig 2.1), if y_t represents the SPM concentration in minute t, this series might be observed over a 6 hours period, t=0,1,2,...,300, starting from 10:30 and ending 16:30, and one may want to forecast the values taken by this series for 17:00 ($h = 25$) and 22:00 $h = 325$ of the same day. As the reasons for the change of SPM concentration are so complex, the series y_t may be considered to be a sequence of random variables. In particular, when standing at time n and contemplating the value that will occur at time $n + h$, one has very little reason to suppose it possible to forecast this value with perfect accuracy, except by incredible luck. Thus y_{n+h} is a random variable, when viewed at time n, and so should be characterised in probabilistic terms. In particular, one could talk about its distribution which will be conditional, given the information available at time n upon which a forecast is based.

Figure 2.3 shows the situation being considered. To fully characterise y_{n+h}, the value to occur at time $n + h$, one needs a complete probability density function, so that statements such as $Pr(0.016 < y_{n+h} < 0.085) = 0.65$ can be made, for any interval. It will generally be quite impossible to completely determine the shape of the density function without making some very strong and highly unreal assumptions about the form of this function. A rather less ambitious procedure is to try to place confidence intervals about the forthcoming value y_{n+h}, so that a statement of the form

$$Pr(B < y_{n+h} \leq A) = 0.95$$

can be made. The points A and B are shown in Fig 2.3 and enable the forecaster to put limits on the value being forecast with a reasonably high degree of confidence of being correct. An example of such a forecast interval is to say, "I believe the SPM concentration will be in the 0.016 g/l - 0.036 g/l range at 22:00 today, with probability 0.95." If your forecasting procedure were a good one and a whole sequence of such forecasts were made, you would expect that the true SPM concentration or whatever would be outside the stated intervals only about 5% of the time. If you can go through life being wrong only 5% of the time, things should turn out very well for you. Such intervals are called interval forecasts. Confidence intervals are sometimes given in practice, although much less often than they should be, but forecasters are usually content with providing just a

a single guess for y_{n+h} that in some way well represents the whole distribution of possible values. An obvious candidate for such a value is an average, such as the mean shown in Fig 2.3 and in this thesis, the word forecast always means the point forecast.

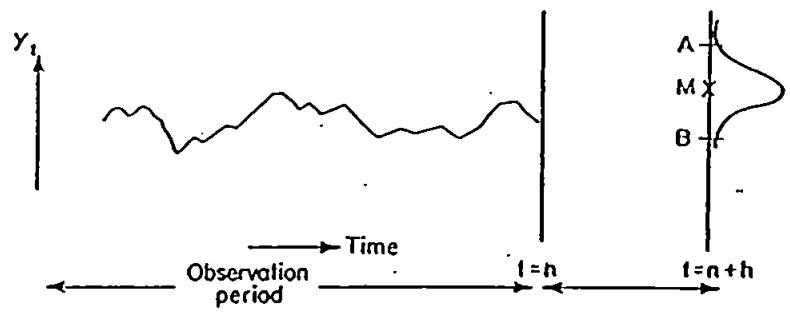


Fig 2.3 Illustration of forecast

2.2 Analysing, Modelling and Forecasting Time Series

2.2.1 Objective of time series modelling

The main reason for setting up a time series model is to enable forecasts of future values to be made. The movements in y_t are explained in terms of its past, or by its position in relation to time. Forecasts are then made by extrapolation.

The statistical approach to forecasting is based on the construction of a model. The model defines a mechanism which is regarded as being capable of having produced the observations in question. Such a model is almost invariably stochastic. If it were used to generate several sets of observations over the same time period, each set of observations would be different, but they would all obey the same probabilistic laws.

Our objective will be to derive models possessing maximum simplicity and the minimum number of parameters consonant with representational adequacy. Obtaining such models

number of parameters consonant with representational adequacy. Obtaining such models is important because:

- They may tell us something about the nature of the system generating the time series;
- They can be used for obtaining optimal forecasts of future values of the series;
- When two or more related time series are under study, the models can be extended to represent dynamic relationships between the series;
- They can be used to derive optimal control policies showing how a variable under one's control should be manipulated so as to minimise disturbances in some dependent variable.

The ability to forecast optimally, to understand dynamic relationships between variables and to control optimality is of great practical importance.
The approach adopted is, first to discuss a class of models which are sufficiently flexible to describe practical situations. In particular, time series are often best represented by nonstationary models in which trends and other pseudo-systematic characteristics which can change with time are treated as statistical rather than as deterministic phenomena. Furthermore, sediment dynamic time series often possess marked seasonal or periodic components themselves capable of change and needing (possibly nonstationary) seasonal statistical models for their description.
The process of model building, which is next discussed, is concerned with relating such a class of statistical models to the data at hand and involves much more than model fitting. Thus identification techniques designed to suggest what particular kind of model might be worth considering, are developed first.
The fitting of the identified model to a time series using recursive least squares or extended least squares method to estimates of the model parameters.
When forecasts are the objective, the fitted statistical model is used directly to generate optimal forecasts by simple recursive calculation. In addition, the fitted model allows one to see exactly how the forecasts utilise past data, to determine the variance of the forecast errors, and to calculate limits within which a future value of the series will lie with a given probability. When the models are extended to represent dynamic relationships, a corresponding iterative cycle of identification, fitting and diagnostic checking is developed

to arrive at the appropriate transfer function. Stochastic models developed earlier are employed in the construction of feed-forward and feedback control schemes.
In this thesis, we shall present methods for building, identifying models for time series and dynamic systems. The methods discussed will be appropriate for discrete (sampled-data) systems, where observation of the system and an opportunity to take control action occur at equally spaced intervals of time. Here we suppose that observations are available at discrete, equi-spatial intervals of time.

2.2.2 Basic Concepts

A time series model is said to be a *univariate* one when only one time series (scalar or vector) and noise series are considered. A time series model is said to be *multivariate* one if more than one time series (scalor or vector) and noise series are considered. A time series is said to be *continuous* when observations are made continuously in time. The term continuous is used for series of this type even when the measured variable can only take a discrete set of values. A time series is said to be *discrete* when observations are taken only at specific times, usually equally spaced. The term discrete is used for series of this type even when the measured variable is a continuous variable as in Figs 2.1 and 2.2.
In this thesis we are mainly concerned with discrete time series. Discrete time series can arise in several ways. Given a continuous time series, we could read off the values at equal intervals of time to give a discrete series called a sampled series. Another type of discrete series occurs when a variable does not have an instantaneous value but we can *aggregate* (or accumulate) the values over equal intervals of time. For example, current velocity and SPM concentration are very important in dealing with sediment transport and observations. A series of T observations will be denoted here by $y_1, ..., y_T$ irrespective of whether they refer to a current velocity or SPM concentration.
There are two aspects to the study of time series analysis and modelling. The aim of analysis is to summarise the properties of a series and to characterise its salient features. This may be done either in the time domain or in the frequency domain. In the time domain, attention is focused on the relationship between observations at different points in time, while in the frequency domain it is cyclical movements which are studied. The two forms of analysis are complementary rather than competitive. The same information is processed in different ways, thereby giving different insights into the nature of the time series. Here, we focus on time series modelling and analysing in the time domain.

It is the modelling of time series as stochastic processes that are primarily of concern here. The sediment dynamic variables such as current velocity and SPM concentration are each taken as a stochastic process and each observation in the stochastic process is a random variable. The observations evolve in time according to certain probabilistic laws. Thus the stochastic process may be defined as a collection of random variables which are ordered in time.

2.3 Some Useful Kinds of Time Series Models

2.3.1 White Noise

Before consideration of how a time series is analysed and forecast, it is necessary to introduce to a few simple but important models, that is methods by which a series can be generated. The simplest possible model gives a purely random series, otherwise known as white noise. This second name is taken from engineering and cannot properly be explained without entering the environment of a method of analysis known as spectral analysis, so no explanation will be attempted. A series is white noise if it has virtually no discernible structure or pattern to it. If such a series is denoted by w_t, for all values of t, the formal definition is that this series is white noise if the sequence $w_t, w_{t-1}, \ldots$ is independent and from a fixed distribution which having mean zero and constant variance σ^2 i.e. $Ew_t = 0$; $Ew_t w_s = \delta_{t,s}\sigma^2$ where $\delta_{t,s} = 1, t = s$ and $\delta_{t,s} = 0, t \neq s$. In the case that w_t is a vector sequence, w_t is a white noise series means that $Ew_t = 0$ and $Ew_t w_s^T = \delta_{t,s}\sigma^2 I$, where 0 is a zero vector and I is an unit matrix.

2.3.2 Backshift Operator and Difference Operator

1. The backshift operator

The backshift operator, z^{-1}, plays an extremely useful role in carrying out algebraic manipulations in time series analysis. It is defined by the transformation

$$z^{-1}y_t = y_{t-1} \tag{2.1}$$

Applying z^{-1} to y_{t-1} yields $z^{-1}y_{t-1} = y_{t-2}$. Substituting from (2.1) gives $z^{-1}(z^{-1}y_t) = z^{-2}y_t = y_{t-2}$ and so, in general,

$$z^{-k}y_t = y_{t-k}, \quad k = 1, 2, 3, ... \tag{2.2}$$

It is logical to complete the definition by letting z^0 have the property $z^0 y_t = y_t$ so that (2.2) holds for all non-negative integers.
The backshift operator can be manipulated in a similar way to any algebraic quantity. Consider a class of so called coloured noise ϵ_t driven by the white noise w_t, which can be represented as following:

$$\epsilon_t = w_t + c_1 w_{t-1} + ... + c_p w_{t-p} \tag{2.3}$$

where $c_1, c_2, ..., c_p$ are constants. The model can also be written in the form:

$$\epsilon_t = (1 + c_1 z^{-1} + c_2 z^{-2} + ... + c_p z^{-p}) w_t \tag{2.4}$$

2. The first difference operator

The first difference operator Δ (sometimes called the forward difference operator), can be manipulated in a similar way to the backshift operator, since $\Delta = 1 - z^{-1}$. The relationship between the two operators can often be usefully exploited. For example,

$$\Delta^2 y_t = (1 - z^{-1})^2 y_t = (1 - 2z^{-1} + z^{-2}) y_t = y_t - 2y_{t-1} + y_{t-2}.$$

The main types of times series model we used in sediment dynamics here are: Autoregressive Model (AR), Autoregressive-moving average Model (ARMA), Autoregressive-Moving Average Exogenous Model (ARMAX) and Multi-input Single Output Model (MISO).

2.3.3 Autoregressive Models (AR Models)

A stochastic model which can be extremely useful in the representation of certain practically occurring series is the so called autoregressive model. In this model, the current value of the process is expressed as a finite, linear aggregate of previous values of the process and a white noise series w_t. Let us denote the values of a process at equally spaced times $t, t-1, t-2, ...$ by $y_t, y_{t-1}, y_{t-2}, ...$, then

$$y_t + a_1 y_{t-1} + a_2 y_{t-2} + ... + a_p y_{t-p} = w_t \tag{2.5}$$

is called an autoregressive (AR) process of order p denoted AR(p). If we define an autoregressive operator of order p by

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + ... + a_p z^{-p}$$

then the autoregressive model can be written economically as

$$A(z^{-1}) y_t = w_t \tag{2.6}$$

In this case, p starting values are required, $x_j, j = 0, 1, ..., p$, and then together with the white noise series w_t the value of y_t are calculated iteratively.
If this model arises and its coefficients are known, it is again easy to use to form forecasts. y_{n+1} will be generated by

$$y_{n+1} = (a_1 y_n + ... + a_p y_{n-p+1}) + w_{n+1}$$

the last term of which is not knowable at time n, so the optimum one-step forecast $f_{n,1}$ is

$$f_{n,1} = a_1 y_n + ... + a_p y_{n-p+1}$$

similarly, y_{n+2} will be generated by

$$y_{n+2} = a_1 y_{n+1} + (a_2 y_n + ... + a_p y_{n-p+2}) + w_{n+2}$$

The last term is not knowable, the term in parentheses is entirely known at time n, and the first term is forecast by $a_1 f_{n,1}$, so the optimum two-step forecast $f_{n,2}$

$$f_{n,2} = a_1 f_{n,1} + (a_2 y_n + ... + a_p y_{n-p+2})$$

It is obvious how further forecasts are formed: One simply writes down the generating mechanism for the value to be forecast, with everything that is known part of the forecast and everything that is not known replaced by its optimum forecast so that we can get the optimum k-step forecast $f_{n,k}, k = 1, 2, ...$ based on the values $y_n, y_{n-1}, ...$

2.3.4 Autoregressive-moving Average Models (ARMA Models)

To achieve greater flexibility in fitting of actual time series, it is sometimes advantageous to include coloured noise series driven by the white noise series. This leads to the autoregressive-moving average model of order p, q denoted ARMA(p, q).

$$y_t + a_1 y_{t-1} + a_2 y_{t-2} + ... + a_p y_{t-p} = w_t + c_1 w_{t-1} + c_2 w_{t-2} + ... + c_r w_{t-r} \tag{2.7}$$

or, set

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_p z^{-p}$$
$$C(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} + \ldots + c_r z^{-r}$$

then the (2.7) is equivalent to the following:

$$A(z^{-1}) y_t = C(z^{-1}) w_t \tag{2.8}$$

A specific example is the ARMA(1,1) generating process

$$y_t = 0.5 y_{t-1} + w_t + 0.3 w_{t-1}$$

so that given a starting value for y_0 and the white noise sequence w_t, y_t is formed iteratively.

Forecasting is straightforward by just using the rules given in the previous two sections, so that in the ARMA(1,1) example y_{n+1} is formed by

$$y_{n+1} = 0.5 y_n + w_{n+1} + 0.3 w_n$$

Then

$$f_{n,1} = 0.5 y_n + 0.3 w_n = 0.5 y_n + 0.3(y_n - f_{n-1,1})$$

by noting that the one-step forecast error is

$$e_{n,1} = w_{n+1}$$

so

$$w_n = e_{n-1,1} = y_n - f_{n-1,1}$$

Further, y_{n+2} is given by

$$y_{n+2} = 0.5 y_{n+1} + w_{n+2} + 0.3 w_{n+1}$$

Both of the last terms are best forecast by their mean values, which are taken as zero,

$$f_{n,2} = 0.5 f_{n,1}$$

and so forth.

Although the ARMA model appears complicated and their statistical properties are difficult to derive, they are of real importance in practice. One reason for this is that there are good theoretical reasons for believing that the ARMA model is the most likely to be found in the real world (Granger, 1980).

2.3.5 Autoregressive Moving Average Exogenous Models (ARMAX Models)

Let y_t and u_t be scalar signals and consider the model structure

$$A(z^{-1})y_t = B(z^{-1})u_t + C(z^{-1})w_t \tag{2.9}$$

where

$$A(z^{-1}) = 1 + a_1 z^{-1} + ... + a_p z^{-p}$$
$$B(z^{-1}) = 1 + b_1 z^{-1} + ... + b_q z^{-q}$$
$$C(z^{-1}) = 1 + c_1 z^{-1} + ... + c_r z^{-r}$$

The model (2.9) can be written explicitly as the difference equation

$$y_t + a_1 y_{t-1} + ... + a_p y_{t-p} = b_1 u_{t-1} + ... + b_q u_{t-q} + w_t + c_1 w_{t-1} + ... + c_r w_{t-r} \tag{2.10}$$

but the form (2.9) using the polynomial formalism will be more convenient. The model (2.9) is called an ARMAX(p, q, r) model, which is short for an ARMA(p, q) model (autoregressive moving average) with an exogenous signal (i.e. an input variable u_t).
Fig 2.4 gives block diagrams of the model (2.9).

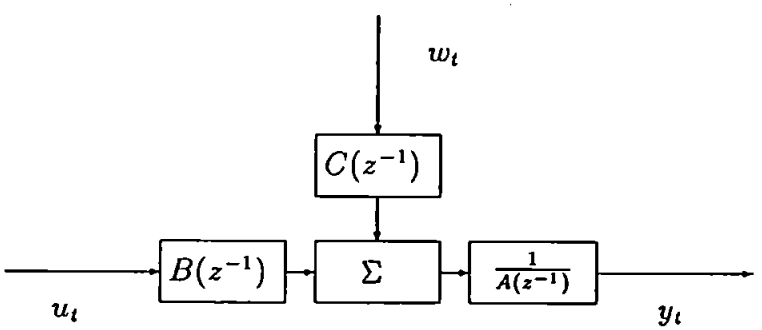


Fig 2.4 Block diagram of an ARMAX model

There are several important special cases of (2.9):

- An autoregressive (AR) model is obtained when $q = r = 0$ (Then a pure time series is modelled as (2.5), i.e. no input signal is assumed to be present.)

- an autoregressive moving average (ARMA) model (2.7) is obtained when $q = 0$. When $A(z^{-1})$ is constrained to certain the factor $1 - z^{-1}$ the model is called autogressive integrated moving average (ARIMA). The model is useful for describing drifting disturbances.

Although the ARMAX model seems to be more complicated than the AR and ARMA model, it can study more than one time series and represent the dynamic relationships between the series. It has a wider application in engineering and economics. The generation process and forecast process are very similar to the AR model and ARMA model.

2.3.6 Multi-Inputs Single-Output Models (MISO Models)

If the system shown in Fig 2.4 contains more than one input variables, we can consider the MISO model as follows:

$$A(z^{-1})y_n = \sum_{i=1}^{s} B_i(z^{-1})u_{i_n} + \epsilon_n \tag{2.11}$$

where $\epsilon_n = C(z^{-1})w_n$ is the system noise and

$$A(z^{-1}) = 1 + A_1 z^{-1} + ... + A_p z^{-p}$$
$$B_i(z^{-1}) = 1 + B_{i_1} z^{-1} + ... + B_{i_{q_i}} z^{-q_i}$$
$$C(z^{-1}) = 1 + C_1 z^{-1} + ... + C_r z^{-r}.$$

The model (2.11) can also written explicitly as the difference equation

$$\begin{aligned} y_n \;=\; & a_1 y_{n-1} + a_2 y_{n-2} + ... + a_p y_{n-p} \\ & + b_{1_1} u_{1_{n-1}} + b_{1_2} u_{1_{n-2}} + ... + b_{1_{q_1}} u_{1_{n-q_1}} \\ & + ... \\ & + b_{s_1} u_{s_{n-1}} + b_{s_2} u_{s_{n-2}} + ... + b_{s_{q_k}} u_{s_{n-q_k}} \\ & + \epsilon_n \end{aligned} \tag{2.12}$$

It is easy to see that if $s = 1$, (2.11) is a ARMAX model. The MISO model can be used in describing and predicting the suspended sediment concentration and the more detail about that will be given in Chapter 6.

Chapter 3

System Identification Theory

3.1 Introduction

Almost without exception, real dynamical systems are subject to random disturbance. In some circumstances such systems can be approximated by deterministic ones by neglecting the random effects. However, to consider them as truly stochastic systems is not only very attractive from a theoretical point of view but it is in fact necessary in order to improve the performance of a system in an engineering context, (Han-fu Chen 1985, Söderström 1989, Hsia 1977).

In order to describe and understand a stochastic system, one first has to construct its mathematical model, which is known as system identification.

System characterisation and system identification are very fundamental problems in system engineering practice. System characterisation is concerned primarily with setting up mathematical models to represent system variable relationships. On the other hand, system identification deals with the choice of a specific model for a class of models which is mathematically equivalent to a given system.

The application of system identification technology goes beyond the boundaries of engineering and physical sciences. Many other fields of study, such as biological sciences, medicine, and economics, can also benefit by employing system identification method to establish quantitative models for the system arising in these areas. Recently, P.C.Young *et al.* (1994) successfully applied the system identification technique to the rainfall-flow dynamical analysis. However, as far as this author knows, there are few applications in

sediment dynamics and sediment transport.
A variety of techniques have been devised over the years for system identification. In general the identification techniques are derived from the optimisation and estimation theories. The purpose of this thesis is to focus on the on-line least squares recursive method as a basic solution to the system identification problem since the least squares method is a classical method frequently practised among scientists in various fields and the other motivation for focusing on the least squares method is that other popular identification methods, such as maximum likelihood, Kalman filtering, instrumental variables and stochastic approximation, can be easily related to the least squares algorithm. Therefore a basis of some degree of integration and unification of many system identification methodologies is introduced.

3.2 Main Concepts of Probability Theory

In this section some basic facts are given from probability theory and random processes. No proofs are given here, they are given in the references (Doob 1953, Lipster and Shiryayev 1977, Loeve 1960, Wang 1965, Chow 1965, Hall 1980 and Han-fu Chen 1985).

3.2.1 Probability Space, Random Variables and Mathematical Expectation

Let $(\Omega, \mathcal{F}, P)$ denote a probability space and ω denote a point of Ω which is also called an elementary event. $\mathcal{F}$ is the $\sigma-$algebra of subset in Ω (i.e., $\mathcal{F}$ has the following properties):

1. $\Omega \in \mathcal{F}$.
2. The complementary set A^c of A belongs to $\mathcal{F}$, if $A \in \mathcal{F}$.
3. $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$, if $A_i \in \mathcal{F}$, $i = 1, 2, ...$

From here it follows immediately that

$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$$

if we notice that $(\bigcap_{i=1}^{\infty} A_i)^c = \bigcup_{i=1}^{\infty} A_i^c$.
A set $A \in \mathcal{F}$ is called a random event. P is called the probability measure on $(\Omega, \mathcal{F})$. It

is a function defined on $\mathcal{F}$ with the properties

1. $P(A) \geq 0, \forall A \in \mathcal{F}$.
2. $P(\Omega) = 1$.
3. $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$, if $A_i \in \mathcal{F}$, and $A_i \bigcap A_j = \emptyset, \forall i \neq j$.

$P(A)$ is called the probability of the random event A.

Let B be any subset of a set $A \in \mathcal{F}$, where A is of probability zero. Then we assume that $B \in \mathcal{F}$ and that it also has probability zero. The probability space with such an extended σ-algebra is called a complete probability space. In the sequel we shall only consider complete probability spaces.

We shall always denote the l-dimensional random Euclidean space by R^l and its Borel σ-algebra by B^l. By a Borel σ-algebra on a topological space we mean the smallest σ-algebra containing all the open sets of the topology. A measurable function $\xi = \xi(\omega)$ defined on $(\Omega, \mathcal{F})$ and valued in (R^l, B^l) is called the l-dimensional random vector.

Let ξ, η be two l-dimensional random vectors. We say that ξ is equal to η with probability one, or almost surely, and denote this by

$$\xi = \eta \quad a.s.$$

if

$$P(\xi \neq \eta) = 0$$

Let ξ be a one-dimensional nonnegative random variable and set

$$A_{ni} = \{\omega : i2^{-n} < \xi \leq (i+1)2^{-n}\}$$

The mathematical expectation $E\xi$ of nonnegative random variable ξ is defined as the integral

$$E\xi = \int_{\Omega} \xi dP = \lim_{n \to \infty} \left[\sum_{i=1}^{n2^n} i2^{-n} P A_{ni} + nP(\xi > n) \right],$$

which may be infinite.

For an arbitrary random variable ξ, define

$$\xi^+ = max(\xi, 0), \quad \xi^- = max(-\xi, 0).$$

These are both nonnegative and hence $E\xi^+$ and $E\xi^-$ are well defined. Notice that

$$\xi = \xi^+ - \xi^-,$$

so it is natural to define

$$E\xi = E\xi^+ - E\xi^-$$

if at least one of $E\xi^+$ and $E\xi^-$ is finite.
If $E|\xi| = E\xi^+ + E\xi^- < \infty$, then ξ is said to be integrable or to have finite expectation.
Let

$$\xi = [\xi^1, ..., \xi^l]^\tau$$

be an l-dimensional random vector. By its distribution function, we mean the function defined by

$$F_\xi(x^1, ..., x^l) = P[\xi^1 < x^1, ..., \xi^l < x^l].$$

If there is a function $f_\xi(x^1, ..., x^l)$ such that

$$F_\xi(x^1, ..., x^l) = \int_{-\infty}^{x^1} ... \int_{-\infty}^{x^l} f_\xi(\lambda^1, ..., \lambda^l) d\lambda^1 ... d\lambda^l,$$

then $f_\xi(x^1, ..., x^l)$ is called the density of the distribution of ξ or simply the density of ξ. When ξ is one-dimensional, then its distribution function and density are denoted by $F_\xi(x)$ and $f_\xi(x)$, respectively. Notice that the mathematical expectation of a random variable ξ can be written as a Lebesgue-Stieltjes integral with respect to its distribution function:

$$E\xi = \int_\Omega \xi dP = \int_{-\infty}^{\infty} x dF_\xi(x).$$

3.2.2 Convergence Theorems

The convergence of a sequence of random variables ξ_n to its limit ξ can take place in several different ways:

1. Convergence with probability one or almost surely means that, with the possible exception of a set of probability zero, for any $\omega \in \Omega$, $\xi_n(\omega) \to \xi(\omega)$, that is,

$$P(\xi_n \to \xi) = 1$$

For this type of convergence, we often write

$$\xi_n \to \xi \quad a.s.$$

2. Convergence in probability means that for any $\epsilon > 0$

$$\lim_{n\to\infty} P\left[|\xi_n - \xi| > \epsilon\right] \to 0$$

and this is denoted by

$$\xi_n \to \xi \quad p.$$

3. Convergence in distribution or weak convergence means that for any x where $F_\xi(x)$ is continuous

$$\lim_{n\to\infty} F_{\xi_n}(x) \to F_\xi(x),$$

and it is denoted by

$$\lim_{n\to\infty} \xi_n = \xi \quad w.$$

4. Convergence in the mean square sense means that

$$\lim_{n\to\infty} E|\xi_n - \xi|^2 = 0.$$

The following diagram explains the relationship of these convergence types

$$\begin{array}{c} \text{convergence a.s.} \Rightarrow \text{convergence p.} \Rightarrow \text{convergence w.} \\ \Uparrow \\ \text{convergence in the mean square sense.} \end{array}$$

Let $E|\xi_n| < \infty$. We now give conditions for $\lim_{n\to\infty} E\xi_n = E\xi$.

Theorem 3.1 (Monotone Convergence Theorem).
If $\xi_n \uparrow \xi, (\xi_n \downarrow \xi)$ *a.s. and* $E\xi_1^- < \infty$ $(E\xi_1^+ < \infty)$, *then* $E\xi_n \uparrow E\xi$ $(E\xi_n \downarrow E\xi)$.

Theorem 3.2 (Fatou Lemma).
If there exists an integrable random variable η *such that* $\eta \le \xi_n$, $(\xi_n \le \eta)$, *then*

$$Einf \lim_{n\to\infty} \xi_n \le inf \lim_{n\to\infty} E\xi_n \quad (sup \lim_{n\to\infty} E\xi_n \le Esup \lim_{n\to\infty} \xi_n).$$

Theorem 3.3 (Dominated Convergence Theorem).
If $\lim_{n\to\infty} \xi_n = \xi$ *a.s. and there exists an integrable random variable* η *such that* $|\xi| \le \eta$, *then*

$$\lim_{n\to\infty} E|\xi_n - \xi| = 0.$$

3.2.3 Independence

Let $A_i \in \mathcal{F}, i = 1, 2, \ldots$. If for any set of indices $\{i_1, \ldots, i_k\}$

$$P\bigcap_{j=1}^{k} A_{i_j} = \prod_{j=1}^{k} PA_{i_j}.$$

then random events are called mutually independent.

We say that $\mathcal{F}_1$ is a sub-σ-algebra of $\mathcal{F}$ if $A \in \mathcal{F}_1$ implies $A \in \mathcal{F}$ and $\mathcal{F}_1$ itself is a σ-algebra of sets in Ω. Sub-σ-algebras $\mathcal{F}_i, i = 1, 2, \ldots$ of $\mathcal{F}$ are called mutually independent if any index set $\{i_1, \ldots, i_k\}$, the random events $A_1, \ldots, A_k$ are mutually independent when $A_j \in \mathcal{F}_{i_j}$ for $j = 1, \ldots, k$.

Let η be an l-dimensional random vector. Denote by $\mathcal{F}^\eta$ the smallest σ-algebra containing all sets of form

$$\{\omega = \eta^{-1}(B), \quad B \in \mathcal{B}^l\}$$

and call it the $\sigma-$algebra generated by η.

Random vectors $\eta_i, i = 1, 2, \ldots$ are called mutually independent if $\sigma-$ algebras $\mathcal{F}^{\eta_i}$ are mutually independent.

If $\{\eta_i\}$ are mutually independent and identically distributed with $E\|\eta_i\| < \infty$, then

$$\lim_{n\to\infty} \quad \frac{1}{n}\sum_{i=1}^{n} \eta_i = E\eta_i \quad a.s.$$

This is called the strong law of large numbers.

Theorem 3.4 (Borel-Cantelli Lemma).

Let $A_1, A_2, \ldots$ be random events.

1. *If $\sum_{i=1}^{\infty} PA_i < \infty$, then $P\bigcap_{i=1}^{\infty}\bigcup_{j=i}^{\infty} A_j = 0$.*
2. *If the events $\{A_i\}$ are mutually independent and $\sum_{i=1}^{\infty} PA_i = \infty$, then $P\bigcap_{i=1}^{\infty}\bigcup_{j=i}^{\infty} A_j = 1$.*

The set $\bigcap_{i=1}^{\infty}\bigcup_{j=i}^{\infty} A_j$ is usually denoted by $\overline{lim}_{j\to\infty} A_j$ and it consists of all ω which appear in an infinite number of A_j.

3.2.4 Conditional Expectation

The relationships between any two quantities in the sequel always permits the failure of that relationship on a set with probability zero. This point will not be mentioned every

time. For example, $\lim_{n\to\infty}\xi_n = \xi$ means convergence with probability one, but we shall frequently omit to write *a.s.*, similarly, the quantifier $\forall\omega \in J$ corresponds to "all ω in J with the possible exception of a set J with zero Probability."
We always denote by I_A the indicator of a set A:

$$I_A = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

and define

$$\int_A \xi dP = E\xi I_A$$

If on $\mathcal{F}$, besides the probability measure P, there is another measure Q such that for $A \in \mathcal{F}, PA = 0$ implies $QA = 0$, then Q is called absolutely continuous with respect to P and this fact is denoted by $Q \ll P$.

Theorem 3.5 (Radon-Nikodym).
If $Q \ll P$, then there exists a nonnegative random variable ξ such that for any $A \in \mathcal{F}$

$$Q(A) = \int_A \xi dP$$

and ξ is defined uniquely in the sense that if there is another nonnegative random variable η with the property

$$Q(A) = \int_A \eta dP \quad \forall A \in \mathcal{F},$$

then $P(\xi \neq \eta) = 0$.

This kind of uniqueness is called to within stochastic equivalence, and ξ is the Radon-Nikodym derivative (or the density of one measure (P) with respect to the other (Q), denote by

$$\xi = \frac{dQ}{dP}.$$

Let $\mathcal{F}_1 \subset \mathcal{F}$ be a sub-σ-algebra and $P^{\mathcal{F}_1}$ be a probability measure on $\mathcal{F}_1$ defined simply by setting

$$P^{\mathcal{F}_1}(A) = P(A). \quad \forall A \in \mathcal{F}_1.$$

Let η be a nonnegative random variable and define

$$Q(A) = \int_A \eta dP = \int_A \eta dP^{\mathcal{F}_1} \quad \forall A \in \mathcal{F}_1. \tag{3.1}$$

Clearly, Q is a measure on $\mathcal{F}_1$ and is absolutely continuous with respect to $P^{\mathcal{F}_1}$, hence by Theorem 3.5 there exists an $\mathcal{F}_1-$measurable nonnegative random variable ξ such that

$$Q(A) = \int_A \xi dP^{\mathcal{F}_1} = \int_A \xi dP. \tag{3.2}$$

Comparing (3.2) with (3.1), we find that there is a nonnegative random variable ξ such that

$$\int_A \eta dP = \int_A \xi dP, \quad \forall A \in \mathcal{F}_1.$$

ξ is called conditional expectation of ξ given $\mathcal{F}_1$ and is denoted by

$$\xi = E(\eta/\mathcal{F}_1) \quad or \quad E^{\mathcal{F}_1}\eta.$$

By Theorem 3.5, $E(\eta/\mathcal{F}_1)$ is unique to within stochastic equivalence.
For the general random variable η (not necessarily nonnegative) if $E\eta$ exists (i.e. at least one of $E\eta^+$ and $E\eta^-$ is finite), then define

$$E(\eta/\mathcal{F}_1) = E(\eta^+/\mathcal{F}_1) - E(\eta^-/\mathcal{F}_1).$$

The conditional expectation $E(\xi/\eta)$ of ξ given a random vector η is defined by

$$E(\xi/\eta) = E(\xi/\mathcal{F}^\eta).$$

Assume $E\|\xi\| < \infty, E\|\eta\| < \infty$. The conditional expectation has the following properties:

1. $E^{\mathcal{F}_1}(a\xi + b\eta) = aE^{\mathcal{F}_1}\xi + bE^{\mathcal{F}_1}\eta$, where a, b, are constants.
2. Let ξ be a random vector. There exists a Borel measurable function $f(\cdot)$ such that

   $$E(\xi/\zeta) = f(\zeta).$$

3. Let $EE^{\mathcal{F}_1}\xi = E\xi$
4. If ξ is $\mathcal{F}_1$-measurable.

   $$E^{\mathcal{F}_1}\xi = \xi \tag{3.3}$$

5. $E^{\mathcal{F}_1}\zeta^\tau\xi = \zeta^\tau E^{\mathcal{F}_1}\xi$, if ξ is $\mathcal{F}_1-$measurable and $E\|\zeta^\tau\xi\| < \infty$.
6. If $\mathcal{F}_1$ and $\mathcal{F}_2$ are sub-σ-algebra with $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}$, then

   $$E^{\mathcal{F}_1}E^{\mathcal{F}_2}\xi = E^{\mathcal{F}_1}\xi.$$

7. If ξ and ζ are independent, then

   $$E(\xi/\zeta) = E\xi. \tag{3.4}$$

8. If $\mathcal{F}_1 = (\Omega, \phi)$, then

$$E^{\mathcal{F}_1}\xi = E\xi$$

The conditional probability $P^{\mathcal{F}_1}A$ or $P(A/\mathcal{F}_1)$ of $A \in \mathcal{F}$ given $\mathcal{F}_1$ is defined by

$$P^{\mathcal{F}_1}(A) = E(I_A/\mathcal{F}_1),$$

and if $\mathcal{F}_1$ is the σ-algebra $\mathcal{F}^\eta$ generated by η, then $P(A/\mathcal{F}^\eta)$ is called the conditional probability of A given η and is denoted by $P^\eta A$ or $P(A/\eta)$. Clearly,

$$P^{\mathcal{F}_1}(A) \geq 0, \quad P^{\mathcal{F}_1}(\Omega) = 1,$$

and

$$P^{\mathcal{F}_1}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P^{\mathcal{F}_1}(A_i)$$

if $A_i \bigcap A_j = \phi, \forall i \neq j$.

Theorems 3.1-3.3 can be extended from expectation to the conditional expectation.

Theorem 3.6.

If $\xi_n \uparrow \xi \quad (\xi_n \downarrow \xi) \quad a.s.$ *and* $E\xi_1^- < \infty \quad (E\xi_1^+ < \infty)$, *then*

$$E^{\mathcal{F}_1}\xi_n \uparrow E^{\mathcal{F}_1}\xi \quad (E^{\mathcal{F}_1}\xi_n \downarrow E^{\mathcal{F}_1}\xi) \quad a.s.$$

Theorem 3.7.

Suppose that η *is an integrable random variable.*

1. If $\eta \leq \xi_n \quad (\xi_n \leq \eta)$, *then*

$$E^{\mathcal{F}_1} inf \lim_{n\to\infty} \xi_n \leq inf \lim_{n\to\infty} E^{\mathcal{F}_1}\xi_n \quad \left(sup \lim_{n\to\infty} E^{\mathcal{F}_1}\xi_n \leq E^{\mathcal{F}_1} sup \lim_{n\to\infty} \xi_n\right) \quad a.s.$$

2. If $|\xi_n| \leq \eta$ *and* $\lim_{n\to\infty} \xi_n = \xi \quad a.s.$, *then*

$$\lim_{n\to\infty} E^{\mathcal{F}_1}|\xi_n - \xi| = 0 \quad a.s.$$

Random vectors ξ and η are called conditionally independent given ζ if

$$P^\zeta(\xi < \mathrm{x}, \eta < \mathrm{y}) = P^\zeta(\xi < \mathrm{x})P^\zeta(\eta < \mathrm{y}), \quad \forall \mathrm{x}, \mathrm{y},$$

where the inequality $\xi < \mathrm{x}$ between vectors should be understood as inequalities between their components.

3.2.5 Stochastic Processes

Let $T = [0, \infty)$ and Let $\mathcal{B}(T)$ be the σ-algebra of Borel sets on T. A function $\xi_t(\omega)$ defined on $(\Omega \times T, \mathcal{F} \times \mathcal{B}(T))$ and taking values in $(R^l, \mathcal{B}^l)$ is called an l-dimensional continuous time stochastic process. If $\xi_t(\omega)$ is only defined at discrete times $t = 0, 1, 2, ...$, then it is called a discrete time (parameter) stochastic process or a random sequence.
For fixed $\omega, \xi_t(\omega)$ is a function of ω and is called a trajectory of the stochastic process.
If for any Borel set B

$$\{(\omega, t): \quad \xi_t(\omega) \in B\} \in \mathcal{F} \times \mathcal{B}(T),$$

then $\xi_t(\omega)$ is called a measurable stochastic process.
We often omit ω and denote a process simply by ξ_t.

Theorem 3.8 (Fubini).
if ξ_t is a measurable stochastic process, then almost all of its trajectories are Borel measurable functions of t. In addition, if $E\xi_t$ exists $\forall t \in T$, then it is also a measurable function. Further, if

$$\int_S E\|\xi_t\| dt < \infty$$

then

$$\int_S \|\xi_t\| dt < \infty \quad a.s.$$

and

$$E \int_S \xi_t dt = \int_S E\xi_t dt$$

where S is any measurable set in T. Two stochastic processes ξ_t and η_t are called stochastically equivalent if

$$P(\xi_t \neq \eta_t) = 0 \quad \forall t \in T,$$

and in this case $\xi_t(\eta_t)$ is called a modification of $\eta_t(\xi_t)$.
If for all $\omega \in \Omega$, with the possible exception of a set of zero probability, the trajectories of ξ_t are continuous (left-continuous or right continuous), then ξ_t is called continuous (left-continuous or right-continuous, respectively) process. A left- or right-continuous process is measurable.
Let $\mathcal{F}_t$ be a family of nondecreasing σ-algebras (i.e., $\mathcal{F}_s \subseteq \mathcal{F}_t, \quad \forall s \leq t$) . If ξ_t is $\mathcal{F}_t$-measurable for any $t \in T$, then we say that ξ_t is $\mathcal{F}_t$-adapted and write $(\xi_t, \mathcal{F}_t)$.

If ξ_t is a measurable process, $E\|\xi_t\| < \infty, \quad \forall t \in T$ and $\{\mathcal{F}_t\}$ is a nondecreasing family of σ-algebra, then in the equivalence class $E(\xi_t/\mathcal{F}_t)$ a modification can be chosen to be $\mathcal{F}_t$ measurable. In the sequel, we always assume that $E(\xi_t/\mathcal{F}_t)$ is so chosen.

3.2.6 Martingales

Definition 3.1.
Let ξ_t be adapted to a nondecreasing family of σ-algabras $\{\mathcal{F}_t\}$ with $E|\xi_t| < \infty$. $(\xi_t, \mathcal{F}_t)$ is called a martingale if

$$E(\xi_t/\mathcal{F}_s) = \xi_s \quad \forall s \le t, \quad s, t \in T$$

Definition 3.2.
Let ξ_t be adapted to a nondecreasing family of σ-algabras $\{\mathcal{F}_t\}$ with $E|\xi_t| < \infty$. $(\xi_t, \mathcal{F}_t)$ is called a supermartingale if

$$E(\xi_t/\mathcal{F}_s) \le \xi_s \quad \forall s \le t, \quad s, t \in T$$

Definition 3.3.
Let ξ_t be adapted to a nondecreasing family of σ-algabras $\{\mathcal{F}_t\}$ with $E|\xi_t| < \infty$. $(\xi_t, \mathcal{F}_t)$ is called a submartingale if

$$E(\xi_t/\mathcal{F}_s) \ge \xi_s \quad \forall s \le t, \quad s, t \in T$$

The above three mentioned definitions also hold for a discrete-time process.

Example 3.1.
Suppose $\eta_i, i = 1, 2, \ldots$ to be a mutually independent random sequence with $E\eta_i = 0, \quad \forall i$. Denote

$$\xi_n = \sum_{i=1}^{n} \eta_i, \quad \mathcal{F}_n = \mathcal{F}_n^{\eta}$$

where $\mathcal{F}_n^{\eta}$ denotes the σ-algebra generated by $\eta_1, \ldots, \eta_n$. We know $E(\xi_m/\mathcal{F}_m) = \xi_m$ by (3.3) and $E(\sum_{i=m+1}^{n} \eta_i/\mathcal{F}_m) = E\sum_{i=m+1}^{n} \eta_i = 0$ by (3.4) for any $m \le n$. Hence

$$E(\xi_n/\mathcal{F}_m) = E(\xi_m + \sum_{i=m+1}^{n} \eta_i/\mathcal{F}_m) = \xi_m$$

and $(\xi_n, \mathcal{F}_n)$ is a martingale.

Theorem 3.9.

Assume $(\xi_n, \mathcal{F}_n)$ *to be a submartingale (supermartingale) and* $\sup_n E\xi_n^+ < \infty \quad (\sup_n E\xi_n^- < \infty)$. *Then* ξ_n *converges to a finite limit* $\xi \quad a.s.$ *as* $n \to \infty$ *and* $E\xi^+ < \infty \quad (E\xi^- < \infty)$).

Corollary 3.1.

If $(\xi_n, \mathcal{F}_n)$*is a nonpositive (nonnegative) submartingale (supermartingale), then* ξ_n *converges to a finite limit as* $n \to \infty$.

Corollary 3.2.

$(\xi_n, \mathcal{F}_n)$ *is a martingale, then* $E|\xi_n| = E\xi_n^+ + E\xi_n^- = 2E\xi_n^+ - E\xi_n = 2E\xi_n^+ - E\xi_1$. *Hence for martingale* $\sup_n E\xi_n^+ < \infty$ *(or* $\sup_n E\xi_n^- < \infty$*) is equivalent to* $\sup_n E|\xi_n| < \infty$.
If $(\xi_n, \mathcal{F}_n)$ *is a martingale, then* $\{x_n\}$ *defined by* $x_1 = \xi_1, ..., x_n = \xi_n - \xi_{n-1}$ *is called a martingale difference sequence. The following two theorems are concerned with local convergence of martingales.*

Theorem 3.10.

Let $x_n = \xi_n - \xi_{n-1}, ..., x_1 = \xi_1$ *be a martingale difference sequence, then* ξ_n *converges a.s. to a finite limit on A, where*

$$A = \left\{ \sum_{i=2}^{\infty} E\left[\frac{(|x_i|^2 I_{[|x_i| \le a_i]} + |x_i|^2 I_{[|x_i| > a_i]})}{\mathcal{F}_{i-1}} \right] < \infty \right\}$$

and a_i *are constants with* $a_i \ge c > 0$.

As a consequence of Theorem 3.10, we obtain:

Theorem 3.11

ξ_n *converges to a finite limit a.s. on A where*

$$A = \left\{ \sum_{i=2}^{\infty} E(\frac{|x_i|^\rho}{\mathcal{F}_{i-1}}) < \infty \right\}, \quad 1 \le \rho \le 2.$$

3.3 Strong Consistency of Least Squares Identification

3.3.1 Introduction

Given a physical system S, in order to predict its evolution, one first has to construct its mathematical model. In some circumstances one can derive it theoretically starting from relationships provided by physics or mechanics; an example is the equations of motion of a satellite in its orbit. However, the mathematical model obtained in such a way may contain a certain number of unknown parameters, for example, the motion equations of a plane derived from the mechanical relationships may include some unknown dynamic coefficients. In many cases one cannot obtain a model of the system from physics and mechanics at all. Consider, for example, the process arising in a complicated chemical reaction. Hence it is of great importance to define the mathematical model for a system based on its inputs and outputs. For example, for an aircraft in flight the change of its rudder angle may be regarded as an input and the three co-ordinates of its position in space may be viewed as the output of the system; for a chemical reaction the product depends on the levels of, say, a temperature, pressure, and a catalytic agent, these can be viewed as system inputs and the product as the system output; for sediment transport process, the suspended sediment concentration depends on the current velocity, pressure, salinity which can be viewed as the inputs of system and suspended sediment concentration can be viewed as the output of the system. The task of system identification is to find the equations of the system. Since the measured data are usually corrupted by random noise, the identified system is a system under random influences. Several aspects must be considered in identification of a stochastic system.

1. *Selection of Model Set* $M(\theta)$. $M(\theta)$ is parameterized by some parameter θ to be selected. The true system S may lie in $M(\theta)$, but for most cases S does not belong to $M(\theta)$. Thus θ has to be chosen such that $M(\theta)$ approximates S as well as possible.

2. *Parameter Estimation.* With $M(\theta)$ having been selected and with input-output data having been obtained, the next step is to construct the estimate $\hat{\theta}$ of θ such that $M(\hat{\theta})$ is

consistent with true data as well as possible.

3. *Properties of the Parameter Estimate.* Having specified the parameter estimate $\hat{\theta}$, one has to determine its properties. It is usually desirable that $\hat{\theta}$, has at least the following property: if $S \in M(\theta)$ then $\hat{\theta}$ asymptotically converges to the true parameter θ as the data increases. This is the so-called consistency problem. In addition, one also wishes to obtain the convergence rate of the estimate, its asymptotic distribution, the efficiency of the estimate, and other related properties.
The consistency problem is discussed here. For the case with the coloured system noise for the least squares identification, both strong consistency and converge rate are given by using the stochastic Lyapunov function series method. (Chen and Ruan, 1987; Chen *et al.* 1996).
Results are presented here are suitable for time series analysis since the dynamic model considered here is nothing but the ARMAX, bilinear and MISO model in the time series analysis.

3.3.2 Review of Convergence Analysis

The strong consistency of parameter estimation has always been one of the main problems in system identification theory. There are many identification algorithms in linear time-invariant stochastic systems which have strong consistency (Ljung *et al.* 1983, Han-fu Chen 1981a, Han-fu Chen 1981b, Han-fu Chen and Guo 1985). Han-fu Chen (1982) has studied the problem for linear time-invariant stochastic systems. A sufficient condition for strong consistency of the least squares identification algorithm has been presented for a white noise model. The convergence rate was also given.

For discrete-time stochastic systems with coloured noise, the strong consistency of the parameter estimates of least squares identification and adaptive control has been studied for various conditions, but the convergence rate of the parameter estimates was not given. The results were extended in Chen and Ruan (1987) for the multivariable input-output ARMAX model. The strong consistency of the coefficient matrix and a better convergence rate were described for conditions weaker than the persistent excitation condition.

In recent years, there has been much study of the identification problem of bilinear sys-

tems. Many practical system models are bilinear, and the realisation of general non-linear problems can be attained by bilinear systems (Krener 1975). Many identification methods, such as recursive least squares, extended least squares, recursive auxiliary variable and recursive prediction error algorithms, have been used in bilinear systems. Simulation studies have been undertaken (Fnaiech and Ljung 1987). The precision of the estimate of parameters is not ideal, since estimates often have large errors. It is necessary to study the conditions needed to guarantee the consistency. The identification algorithms discussed in Ahmed (1986) and Wang and Lu (1987) have yielded good simulation results, but they did not give conditions for consistency and theoretical proofs.

The strong consistency of the coefficient matrix and the noise covariance matrix were given in Zhang (1983), but the model class considered did not include the multiplicator of input and output. The single-input single-output non-linear system is more general than bilinear systems, but its noise model is linear and does not include the multiplicator of noise and input. Although the analysis indicated the convergence analysis of least squares identification, the convergence rate was not studied. Also the restrictions to the noise series and persistent excitation are rather strict requirements.

In Chen *et al.* (1996), identification problems for a class of discrete-time bilinear stochastic systems are discussed. We remove certain strict requirements and do not need the noise series to be stationary or quasi-stationary. Also the conditional expectation of the variance of the noise series is allowed to be unbounded for all stochastic variables and time. A simple condition is presented to guarantee the least square identification to have strong consistency in the case of systems with coloured noise. In order not to make the problem too complicated, we consider here only the single-input single-output case. The results are easily extended to the multivariable input-output case or more generally to the ARMAX model (Billings and Voon 1984) and the non-linear model with linear parameters. We consider a condition which is weaker than the persistent excitation condition given by Han-fu Chen and Guo (1985) and Ljung and Söderström (1983) for white and coloured noise. The strong consistency of the parameter estimate using extended least squares, and the convergence rate are proved.

3.3.3 Least Squares Theory

Least squares theory was first proposed by Karl Gauss for carrying out his work in orbit prediction of planets. Least squares theory has since become a major tool for parameter estimation from experimental data. Although there are several other estimation methods available, such as maximum likelihood, Bayes method and so on, the least squares method continues to be the most well known among engineers and scientists. The reason for its popularity is that the method is easier to comprehend than others and does not require a knowledge of mathematical statistics. Furthermore, the least squares method may provide solutions even in cases when other methods have failed.

Estimates obtained by the least squares method also have optimal statistical properties: they are consistent, unbiased and efficient. It also turns out that many estimation algorithms that are used for system identification can be interpreted as least squares procedures. Therefore, it is possible to unify many identification techniques in the framework of least squares theory.

The least squares technique provides us with a mathematical procedure by which a model can achieve a best fit to experimental data in the sense of minimum-error-squares. Suppose there is a variable y that is related linearly to a set of variables $\mathbf{x} = (x_1, x_2, ..., x_n)$, that is

$$y = \theta_1 x_1 + \theta_2 x_2 + ... + \theta_n x_n \tag{3.5}$$

in which $\theta = (\theta_1, \theta_2, ..., \theta_n)$ is a set of constant parameters. We assume here that θ_i, are unknown and we wish to estimate their values by observing the variables y and $\mathbf{x}$ at different times.

Let us assume that a sequence of m observations on both y and $\mathbf{x}$ has been made at times $t_1, t_2, ..., t_m$, and we denote the measured data by $y(i)$ and $x_1(i), ..., x_m(i), i = 1, ..., m$. Now we can relate these data by the following set of m linear equations:

$$y(i) = \theta_1 x_1(i) + \theta_2 x_2(i) + ... + \theta_n x_n(i), \quad i = 1, 2, .., m \tag{3.6}$$

In statistical literature, equation (3.6) is called a regression function, and θ_i are the regression coefficients.

The system of equation (3.6) can be conveniently arranged into a simple matrix form

$$\mathbf{y} = \mathbf{X}\theta \tag{3.7}$$

where

$$\mathrm{y} = \begin{pmatrix} y(1) \\ y(2) \\ \vdots \\ y(m) \end{pmatrix}, \qquad \mathrm{X} = \begin{pmatrix} x_1(1) & \dots & x_n(1) \\ x_1(2) & \dots & x_n(2) \\ \vdots & \dots & \vdots \\ x_1(m) & \dots & x_n(m) \end{pmatrix}, \qquad \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{pmatrix},$$

To be able to estimate the n parameters θ_i, it is necessary that $m \geq n$. If $m = n$, then we can solve θ uniquely from equation (3.7) by

$$\hat{\theta} = \mathrm{X}^{-1}\mathrm{y} \tag{3.8}$$

provided that X^{-1}, the inverse of the square matrix X, exists. $\hat{\theta}$ denotes the estimate of θ. However, when $m > n$, it is generally not possible to determine a set of θ_i exactly satisfying all m equations (3.6) because the data may be complicated by random measurement noise, error in the model, or a combination of both. The alternative then is to determine θ on the basis of least-error-squares.

Define an error vector $\epsilon = (\epsilon_1, \dots, \epsilon_m)^\tau$ and let

$$\epsilon = \mathrm{y} - \mathrm{X}\theta \tag{3.9}$$

Now we will choose $\hat{\theta}$ in such a way that the criterion J

$$\mathrm{J} = \sum_{i=1}^{m} \epsilon_i^2 = \epsilon^\tau \epsilon \tag{3.10}$$

is minimised. To carry out the minimisation, we express

$$\begin{aligned} \mathrm{J} &= (\mathrm{y} - \mathrm{X}\theta)^\tau(\mathrm{y} - \mathrm{X}\theta) \\ &= \mathrm{y}^\tau\mathrm{y} - \theta^\tau\mathrm{X}^\tau\mathrm{y} - \mathrm{y}^\tau\mathrm{X}\theta + \theta^\tau\mathrm{X}^\tau\mathrm{X}\theta \end{aligned}$$

Differentiate J with respect to θ and equate the result to zero to determine the conditions on the estimate $\hat{\theta}$ that minimises J. Thus

$$\frac{\partial J}{\partial \theta}\Big|_{\theta=\hat{\theta}} = -2\mathrm{X}^\tau\mathrm{y} + 2\mathrm{X}^\tau\mathrm{X}\hat{\theta} = 0$$

This yields

$$\mathrm{X}^\tau\mathrm{X}\hat{\theta} = \mathrm{X}^\tau\mathrm{y} \tag{3.11}$$

from which $\hat{\theta}$ can be solved as

$$\hat{\theta} = (\mathrm{X}^\tau\mathrm{X})^{-1}\mathrm{X}^\tau\mathrm{y} \tag{3.12}$$

This result is called the least squares estimator (LSE) of θ. Equation (3.11) is referred to as the normal equation and ϵ is called the residual in statistical literature.

The above result is derived based on a criterion J that weights every error ϵ_i equally. We often refer to this result as ordinary least squares. This formulation can be generalised, however, to allow each error term to be weighted differently. Let $\mathbf{W}$ be the desired weighting matrix. Then the weighted error criterion becomes

$$\begin{aligned} \mathbf{J_w} &= \epsilon^T \mathbf{W} \epsilon \\ &= (\mathbf{y} - \mathbf{X}\theta)^T \mathbf{W} (\mathbf{y} - \mathbf{X}\theta) \end{aligned}$$

Here $\mathbf{W}$ is restricted to being a symmetric positive definite matrix. Minimisation of $\mathbf{J_w}$ with respect to θ yields the weighted least squares estimator (WLSE) of $\hat{\theta}_\mathbf{w}$:

$$\hat{\theta}_\mathbf{w} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{y} \tag{3.13}$$

It is easy to see that when $\mathbf{W}$ is chosen as an identity matrix $\mathbf{I}$, $\hat{\theta}_\mathbf{w}$ reduces to $\hat{\theta}$.

1. Statistical Properties of Least Squares Estimators

In here, we examine the qualities of the least squares estimators derived above. To facilitate the discussion, we wish to focus on the model equation (3.7) in which the vector ϵ is included to account for the measurement noise or model error. Thus we have the noise-disturbed system equation

$$\mathbf{y} = \mathbf{X}\theta + \epsilon \tag{3.14}$$

We assume here that ϵ is a stationary random vector with zero mean value, that is, $E[\epsilon] = 0$. Furthermore, ϵ is uncorrelated with $\mathbf{y}$ and $\mathbf{X}$. Based on these assumptions about ϵ, we wish to know just how good, or how accurate, are the parameter estimates given by equation (3.12) and (3.13).

In general, $\hat{\theta}$ is a random variable. Its accuracy can be conveniently measured by a number of statistical properties such as bias, error covariance, efficiency, and consistency. First we show that $\hat{\theta}$ is unbiased, meaning that $E\hat{\theta} = \theta$. Substituting equation (3.12) into equation (3.14), we have

$$\hat{\theta} = \theta + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \epsilon \tag{3.15}$$

Taking the expectation on both sides of the equation (3.15) and applying the property $E[\epsilon] = 0$, we obtain the desired result

$$E[\hat{\theta}] = \mathrm{E}[\theta] + \mathrm{E}[(\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T]\mathrm{E}[\epsilon] = \theta \tag{3.16}$$

The covariance matrix corresponding to the estimate error $\hat{\theta} - \theta$ is

$$\begin{aligned}\Psi &= E\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T\} \\ &= E\{[(\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\epsilon][(\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\epsilon]^T\} \\ &= (\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\mathrm{E}\{\epsilon\epsilon^T\}\mathrm{X}(\mathrm{X}^T\mathrm{X})^{-1}.\end{aligned}$$

Define the covariance matrix of the error vector ϵ to be

$$\mathrm{R} = \mathrm{E}[\epsilon\epsilon^T], \tag{3.17}$$

Ψ is reduced to

$$\Psi = (\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\mathrm{R}\mathrm{X}(\mathrm{X}^T\mathrm{X})^{-1}. \tag{3.18}$$

Following the same procedure, we can also show that the error covariance of $\hat{\theta} - \theta$ is

$$\Psi_{\mathrm{W}} = (\mathrm{X}^T\mathrm{W}\mathrm{X})^{-1}\mathrm{X}^T\mathrm{W}\mathrm{R}\mathrm{W}^T\mathrm{X}(\mathrm{X}^T\mathrm{W}\mathrm{X})^{-1}. \tag{3.19}$$

At this point, it is interesting to point out that Ψ_{W} can be greatly simplified if we let the weighting matrix W be $\mathrm{W} = \mathrm{R}^{-1}$,

$$\Psi_{\mathrm{W}}(\mathrm{W} = \mathrm{R}^{-1}) = (\mathrm{X}^T\mathrm{R}^{-1}\mathrm{X})^{-1} = \Psi_{\mathrm{MV}}. \tag{3.20}$$

The corresponding estimator $\hat{\theta}_{\mathrm{W}}$ is

$$\hat{\theta}_{\mathrm{W}}(\mathrm{W} = \mathrm{R}^{-1}) = (\mathrm{X}^T\mathrm{R}^{-1}\mathrm{X})^{-1}\mathrm{X}^T\mathrm{R}^{-1}\mathrm{y} = \hat{\theta}_{\mathrm{MV}}. \tag{3.21}$$

The error covariance Ψ_{MV} in equation (3.20) has a very important property: that is, Ψ_{MV} is a minimum error covariance matrix in the sense that for any other choice of weighting matrices W

$$\Psi_{\mathrm{MV}} \leq \Psi_{\mathrm{W}} \tag{3.22}$$

By definition, a positive definite matrix Ψ_{MV} is less than or equal to Ψ_{W} if the difference $\Psi_{\mathrm{MV}} - \Psi_{\mathrm{W}}$ is non-negative definite. The subscript MV in Ψ_{MV} and $\hat{\theta}_{\mathrm{MW}}$ denotes the minimum variance property. The proof of $\Psi_{\mathrm{MV}} \leq \Psi_{\mathrm{W}}$ is somewhat involved, and interested readers can see Deutsch (1965). The estimator $\hat{\theta}_{\mathrm{MV}}$ in equation (3.21) is called the minimum variance estimator, or Markov estimator. Thus we see that θ_{MV} is the best linear unbiased estimator.

Now let us examine another interesting case. When the noise $\epsilon(\mathrm{i}), I = 1, 2, ...,$ are identically distributed and independent with zero mean and variance σ^2, the covariance $\mathbf{R}$ becomes

$$R = E[\epsilon\epsilon^{\mathrm{T}}] = \sigma^2\mathbf{I} \tag{3.23}$$

In this case, both Ψ and Ψ_{MV} are identical:

$$\Psi = \Psi_{\mathrm{MV}} = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1} \tag{3.24}$$

This implies that the corresponding $LSE\ \hat{\theta}$ is a minimum variance estimator. $\hat{\theta}$ is called an efficient estimator.

Last, we wish to show that the $LSE\ \hat{\theta}$ is also a consistent estimator. Rewrite the error covariance matrix Ψ in the form (assume $\mathbf{R} = \sigma^2\mathbf{I}$)

$$\Psi = \sigma^2\left((\mathbf{X}^T\mathbf{X})^{-1}\right) = \frac{\sigma^2}{m}\left(\frac{1}{m}(\mathbf{X}^T\mathbf{X})\right)^{-1} \tag{3.25}$$

in which m is the number of equations in the vector equation (3.14). Assume that $\lim_{m\to\infty}[(1/m)\mathbf{X}^T\mathbf{X}]^{-1} = \Gamma$, where Γ is a non-singular constant matrix. Then

$$\lim_{m\to\infty}\Psi = \lim_{m\to\infty}\frac{\sigma^2}{m}\left(\frac{1}{m}(\mathbf{X}^T\mathbf{X})\right)^{-1} = 0 \tag{3.26}$$

Zero error covariance means that $\hat{\theta} = \theta$ at $m \to \infty$. This convergence property indicates that $\hat{\theta}$ is a consistent estimator.

We have shown that the LSE in the presence of white noise is unbiased efficient, and consistent. Finally, we wish to note that the $LSE\ \hat{\theta}$ is also identical to the maximum likelihood estimator (MLE) when the noise ϵ is Gaussian-distributed. This important property is examined in (Hsia, 1977). Thus we see that the least squares technique does indeed have many advantages.

2. Recursive Least Squares Estimation

We derive a recursive algorithm from the basic least squares solution in equation (3.14). The need for a recursive solution arises when fresh data continuously in supply and we wish to improve our parameter estimates by making use of this new information. With a recursive formula, the estimates can be updated step by step without repeatedly computing the matrix solution of equation (3.7), in which the matrix inversion is quite time-consuming. This recursive solution procedure is often referred as on-line identification.

Recall that the vector equation (3.7) consists of a set of m equations. Let us introduce m as a subscript to y and X in equation (3.7). We have

$$\mathrm{y_m} = \mathrm{X_m}\theta \tag{3.27}$$

Furthermore, denote $\hat{\theta}$ in equation (3.12) as $\hat{\theta}(\mathrm{m})$

$$\hat{\theta}(\mathrm{m}) = (\mathrm{X_m^\tau X_m})^{-1}\mathrm{X_m^\tau y_m} \tag{3.28}$$

Suppose we have obtained a new equation, the $(m+1)$th, as

$$y(m+1) = \theta_1 x_1(m+1) + \theta_2 x_2(m+1) + ... + \theta_n x_n(m+1). \tag{3.29}$$

Define

$$\mathrm{x^\tau(m+1)} = [x_1(m+1), x_2(m+1), ..., x_n(m+1)]. \tag{3.30}$$

We then have

$$y(m+1) = x^\tau(m+1)\theta \tag{3.31}$$

Now the system of $m+1$ equations can be written as

$$y_{m+1} = \mathrm{X_{m+1}^\tau}\theta \tag{3.32}$$

in which

$$\mathrm{y}_{m+1} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(m) \\ - - - - \\ y(m+1) \end{bmatrix} = \begin{bmatrix} \mathrm{y}_m \\ \\ - - - - \\ y(m+1) \end{bmatrix} \tag{3.33}$$

$$y_{m+1} = \mathrm{X_{m+1}^\tau}\theta \tag{3.34}$$

$$\mathrm{X}_{m+1} = \begin{bmatrix} x_1(1) & \dots & x_n(1) \\ \vdots & \dots & \vdots \\ x_1(m) & \dots & x_n(m) \\ - - - - & - - - - & - - - \\ x_1(m+1) & \dots & x_n(m+1) \end{bmatrix} = \begin{bmatrix} \mathrm{X}_m \\ \\ - - - - \\ \mathrm{x^\tau}(m+1) \end{bmatrix} \tag{3.35}$$

The new least squares estimator is

$$\hat{\theta}(m+1) = (\mathrm{X_{m+1}^\tau X_{m+1}})^{-1}\mathrm{X_{m+1}^\tau y_{m+1}} \tag{3.36}$$

t is apparent that to obtain $\hat{\theta}(m+1)$, we must invert an $n \times n$ matrix. The obvious question here is whether or not we can calculate $\hat{\theta}(m+1)$ by simply updating the previous estimate $\hat{\theta}(m)$ without matrix inversion. The answer is yes, and we derive the updating algorithm below.
The following well-known matrix inversion lemma is introduced as follows:

Lemma 3.1 (Åström, 1968)
Let **A**,**C**, *and* **A**+**BCD** *be non-singular square matrices; then the following matrix identity holds:*

$$(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1}+\mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1} \tag{3.37}$$

Define the matrix $\mathbf{P}(m)$ as:

$$\mathbf{P}(m) = (\mathbf{X}_{\mathrm{m}}^{\tau}\mathbf{X}_{\mathrm{m}})^{-1} \tag{3.38}$$

Therefore

$$\mathbf{P}(m+1) = (\mathbf{X}_{\mathrm{m+1}}^{\tau}\mathbf{X}_{\mathrm{m+1}})^{-1}$$

Substituting equation (3.35) and applying the matrix inversion lemma, $\mathbf{P}(m+1)$ can be rewritten as follows:

$$\begin{aligned}
\mathbf{P}(m+1) &= [\mathbf{P}(m)^{-1} + \mathbf{x}(m+1)\mathbf{x}^{\tau}(m+1)]^{-1} \\
&= \mathbf{P}(m) - \mathbf{P}(m)\mathbf{x}(m+1) \\
&\quad \times [1 + \mathbf{x}^{\tau}(m+1)\mathbf{P}(m)\mathbf{x}(m+1)]^{-1}\mathbf{x}^{\tau}(m+1)\mathbf{P}(m).
\end{aligned} \tag{3.39}$$

In view of equation (3.36), we can see that

$$\begin{aligned}
\hat{\theta}(m+1) &= \mathbf{P}(m+1)[\mathbf{X}_m^{\tau}\mathbf{y}_m + \mathbf{x}(m+1)\mathbf{y}(m+1)] \\
&= \mathbf{P}(m)\mathbf{X}_{\mathrm{m}}^{\tau}\mathbf{y}_{\mathrm{m}} - \mathbf{P}(m)\mathbf{x}(m+1)[1 + \mathbf{x}^{\tau}(m+1) \\
&\quad \times \mathbf{P}(m)\mathbf{x}(m+1)]^{-1}\mathbf{x}^{\tau}(m+1)\mathbf{P}(m)\mathbf{X}_m^{\tau}\mathbf{y}_m \\
&\quad + \mathbf{P}(m)\mathbf{x}(m+1)\mathbf{y}(m+1) - \mathbf{P}(m)\mathbf{x}(m+1) \\
&\quad \times [1 + \mathbf{x}^{\tau}(m+1)\mathbf{P}(m)\mathbf{x}(m+1)]^{-1} \\
&\quad \times \mathbf{x}^{\tau}(m+1)\mathbf{P}(m)\mathbf{x}(m+1)]^{-1}\mathbf{y}(m+1)
\end{aligned} \tag{3.40}$$

We can rearrange the last two terms in the form of

$$\begin{aligned}
&\mathbf{P}(m)\mathbf{x}(m+1)[1 + \mathbf{x}^{\tau}(m+1)\mathbf{P}(m)\mathbf{x}(m+1)]^{-1} \\
&\quad \times [1 + \mathbf{x}^{\tau}(m+1)\mathbf{P}(m)\mathbf{x}(m+1) - \mathbf{x}^{\tau}(m+1)\mathbf{P}(m)\mathbf{x}(m+1)]\mathbf{y}(m+1) \\
&\qquad = \mathbf{P}(m)\mathbf{x}(m+1)[1 + \mathbf{x}^{\tau}(m+1)\mathbf{P}(m)\mathbf{x}(m+1)]^{-1}\mathbf{y}(m+1)
\end{aligned} \tag{3.41}$$

But we recognise from equations (3.28) and (3.36) that

$$\hat{\theta}(m) = \mathbf{P}(m)\mathbf{X}_m^T\mathbf{y}_m$$

Thus $\hat{\theta}(m+1)$ can finally be simplified to the form

$$\begin{aligned}\hat{\theta}(m+1) &= \hat{\theta}(m) + \mathbf{P}(m)\mathbf{x}(m+1)[1 + \mathbf{x}^T(m+1)\mathbf{P}(m)\mathbf{x}(m+1)]^{-1} \cdot \\ &\quad \times[\mathbf{y}(m+1) - \mathbf{x}^T(m+1)\hat{\theta}(m)] \end{aligned} \tag{3.42}$$

The result above simply shows that the new estimate is given by the old estimate plus a correction term. The matrix $\mathbf{P}(m)$ in the correction term can be updated by the recursive formula in equation (3.39). It is clear that in both formulas we have completely eliminated the necessity of matrix inversion (we note that the term $[1 + \mathbf{x}^T(m+1)\mathbf{P}(m)\mathbf{x}(m+1)]$ is a scalar) and therefore that the computational efficiency is dramatically improved for updating the estimate $\hat{\theta}$.

The recursive equation (3.42) has a very strong intuitive appeal. We notice that the correction term is proportional to the quantity $\mathbf{y}(m+1) - \mathbf{x}^T(m+1)\hat{\theta}(m)$, which represents the error of fitting the previous estimate $\hat{\theta}(m)$ to the new data $\mathbf{y}(m+1)$ and $\mathbf{x}^T(m+1)$. The vector $\mathbf{P}(m)\mathbf{x}(m+1)[1 + \mathbf{x}^T(m+1)\mathbf{P}(m)\mathbf{x}(m+1)]^{-1}$ determines how the fitting error is weighted in the correction of $\hat{\theta}(m)$. Another interesting fact is that $\mathbf{P}(m)$ can be related to the error covariance matrix Ψ defined by equation (3.18). It shows that $\mathbf{P}(m) = \Psi(m)/\sigma^2$, which means that $\mathbf{P}(m)$ is a direct measure of the error covariance at each m. As we have shown in equations (3.26) and (3.38), $\mathbf{P}(m) = \mathbf{0}$ at the limit $m \to \infty$. We have shown in this section that recursive least squares estimation can be easily carried out by the following recursive algorithm:

$$\hat{\theta}(m+1) = \hat{\theta}(m) + \gamma(m+1)\mathbf{P}(m)\mathbf{x}(m+1)[\mathbf{y}(m+1) - \mathbf{x}^T(m+1)\hat{\theta}(m)] \tag{3.43}$$

$$\mathbf{P}(m+1) = \mathbf{P}(m) - \gamma(m+1)\mathbf{P}(m)\mathbf{x}(m+1)\mathbf{x}^T(m+1)\mathbf{P}(m) \tag{3.44}$$

where

$$\gamma(m+1) = 1/[1 + \mathbf{P}(m)\mathbf{x}(m+1)\mathbf{x}^T(m+1)\mathbf{P}(m)].$$

Therefore, by starting with an initial estimate $\hat{\theta}(0)$ and the corresponding $\mathbf{P}(0)$, we can recursively update $\hat{\theta}$ while new observations are continuously obtained.

3.3.4 Extended Least Squares Identification and Its Convergence Analysis of ARMAX Model

1. Statement of the problem

We assume that the system to be identified is a discrete multivariable inputs and multivariable outputs (MIMO) ARMAX Model and that it can be represented by

$$A(z^{-1})y_n = B(z^{-1})u_n + \epsilon_n \tag{3.45}$$

where $\epsilon_n = C(z^{-1})w_n$ is the system noise and

$$\begin{aligned} A(z^{-1}) &= I_m + A_1 z^{-1} + ... + A_p z^{-p}, \\ B(z^{-1}) &= B_1 z^{-1} + B_2 z^{-2} + ... + B_p z^{-q}, \\ C(z^{-1}) &= I_m + C_1 z^{-1} + ... + C_r z^{-r}. \end{aligned} \tag{3.46}$$

and y_n, u_n, w_n are $m-$, $l-$ and $m-$ dimensional vectors respectively and where z^{-1} is a unit delay operator, A_i, B_j, C_k are $m \times m, m \times 1, m \times m$ unknown matrices to be estimated. I_m is an $m \times m$ unit matrix.

Set

$$\theta^\tau = [-A_1, \ldots, -A_p, B_1, \ldots, B_q, C_1, \ldots, C_r] \tag{3.47}$$

$$x_n^\tau = [y_{n-1}^\tau, \ldots, y_{n-p}^\tau, u_{n-1}^\tau, \ldots, u_{n-q}^\tau, e_{n-1}^\tau, \ldots, e_{n-r}^\tau] \tag{3.48}$$

$$e_n = y_n - \theta_n^\tau x_n \tag{3.49}$$

where θ_n is the estimate of θ at time n.

It is easy to see that the relation (3.45) can also be written as

$$y_n = \theta^\tau x_n + C(z^{-1})w_n + e_n - C(z^{-1})e_n \tag{3.50}$$

Two representations (3.45) and (3.50) of the system are equivalent but can be used differently.

Definition 3.4

A d- dimensional vector $x = (x_1, \ldots, x_d)^\tau$, the mode of x is defined as $\|x\| = (\sum_{i=1}^d x_i^2)^{\frac{1}{2}} = (x^\tau x)^{\frac{1}{2}}$.

Definition 3.5

An $m \times n$ matrix A, the mode of A is defined as $\|A\| = (\lambda_{max}(A^\tau A))^{\frac{1}{2}}$.

where $\lambda_{max}(\cdot)$ denotes the maximum eigenvalue of the matrix.
We define

$$r_n = r_{n-1} + \|x_n\|^2, \quad r_0 = 1, \quad i.e., \quad r_n = 1 + \sum_{i=1}^{n} \|x_i\|^2 \tag{3.51}$$

and

$$\mathcal{F}_n = \sigma\{w_i : \quad i \leq n\}. \tag{3.52}$$

which is the σ-algebra generated by $\{w_i : \quad i \leq n\}$.
We assume that the system noise ϵ_n is driven by a martingale difference sequence $\{w_n\}$ that is,

$$w_n = 0, \quad n \leq 0,$$
$$E(w_n/\mathcal{F}_{n-1}) = 0, \quad E(w^\tau w_n/\mathcal{F}_{n-1}) \leq \xi_0 r_{n-1}^h, \quad n > 0; \quad 0 \leq h < 1. \tag{3.53}$$

where ξ_0 ($\xi_0 > 0 \quad a.s.$) denotes a random variable which depends on ω but is independent of n, h is a constant and

$$k_0 = E\xi_0 < \infty. \tag{3.54}$$

where k_0 is a constant. We assume that $E(\|u_n\|^2) < \infty$.

2. The extended least squares algorithm

In order to identify the system parameter matrix θ, we make use of the following recursive algorithms:

$$\theta_{n+1} = \theta_n + K_{n+1}(y_{n+1}^\tau - x_{n+1}^\tau \theta_n), \tag{3.55}$$

$$K_{n+1} = R_n x_{n+1}/(1 + x_{n+1}^\tau R_n x_{n+1}), \tag{3.56}$$

$$R_{n+1} = (I_d - K_{n+1} x_{n+1}^\tau) R_n \tag{3.57}$$

where x_1 and θ_0 are any deterministic vector and matrix respectively, $R_0 = dI_d$, I_d is a $d \times d$ unit matrix and $d = mp + lq + mr$.

The recursive algorithm (3.55)-(3.57) are called the extended least squares method (ELSM) when the regression vector x_n is determined by (3.48)-(3.49).

3. Convergence Theorem

The convergence theorem are given as follows and the proofs of them are given in Chen *et al.* (1987).

Theorem 3.12

For the system and the algorithm defined above, let the following conditions **T1a**, **T1b** *and* **T1c** *be satisfied:*

(T1a) The transfer matrix $C^{-1}(z^{-1}) - I_m/2$ is strictly positive.
(i.e., zeroes of det C(z) are outside the closed unit disk and

$$C^{-1}exp(i\omega) + C^{-1}exp(-i\omega) - I_m > 0, \quad \omega \in [0, 2\pi].)$$

(T1b) There exists a constant $k_2 > 1$, such that

$$\lim_{n\to\infty} r_n^h (logr_n)^{k_2}/\lambda_{min}^n = 0$$

where $\lambda_{max}^n, \lambda_{min}^n$ denote respectively the maximum and minimum eigenvalues of the matrix R_n^{-1}, where h is given by (3.53).

(T1c) $\lim_{n-\infty} r_n = \infty \qquad a.s.$

Then

I) $\lim_{n-\infty} \theta_n = \theta \qquad a.s.$

II) $\|\tilde{\theta_n}\| = o(r_n^h (logr_n)^{k_2}/\lambda_{min}^n)^{1/2}.$

The proof of Theorem 3.12 see Chen *et al.* (1987) and more results of special cases are given as the following corollaries:

Corollary 3.3.

If the condition **T1c** *of Theorem 3.12 is not satisfied, then the conclusions of Theorem*

3.12 should be modified as follows:

I) $\lim_{n\to\infty}\theta_n = \theta \quad a.s. \quad on \quad \{\omega : r_n \to \infty\}.$

II) $\|\tilde{\theta_n}\| = o(r_n^h(logr_n)^{k_2}/\lambda_{min}^n)^{1/2}. \quad a.s. \quad on \quad \{\omega : r_n \to \infty\}.$

Corollary 3.4.

If the condition **T1b** *of Theorem 3.12 is replaced by*

$$\lambda_{max}^n/\lambda_{min}^n \leq \xi_2$$

where ξ_2 ($\xi_2 > 0$ *a.s.*) *is a random variable and* $E\xi_2 < \infty$ *, then*

I) $\lim_{n\to\infty}\theta_n = \theta \quad a.s.$

II) $\|\tilde{\theta_n}\| = o(r_n^{h-1}(logr_n)^{k_2})^{1/2}$

Corollary 3.5.

If the restriction on $\{w_i\}$ *given by (3.53) is replaced by the following conditions, i.e.,*

$$w_n = 0,\ n \leq 0; \quad E(w_n/\mathcal{F}_{n-1}) = 0,\ E(w_n^T w_n/\mathcal{F}_{n-1}) \leq k_0,\ n > 0,$$

Then

(II) $\lim_{n\to\infty}\theta_n = \theta \quad a.s.$

II) $\|\tilde{\theta_n}\| = o((logr_n)^{k_2}/\lambda_{min}^n)^{1/2}$

3.3.5 Least Squares Identification and Its Convergence Analysis of Bilinear Time Series Model

1. Statement of the problem

The special types of bilinear system are considered in here.

$$A(z^{-1})y_n = B(z^{-1})u_{n-d} + C(z^{-1})w_n + \sum_{i=1}^{l}\sum_{j=0}^{m} a_{ij}y_{n-i}u_{n-j-d} \tag{3.58}$$

where z^{-1} is the unit backshift operator, $d \leq 1$ is a time delay, $\{w_n\}$ is a noise sequence, $A(z^{-1})$, $B(z^{-1})$ and $C(z^{-1})$ are the polynomials

$$A(z^{-1}) \ = \ 1 - a_1 z^{-1} - \ldots - a_p z^{-p}$$

$$
\begin{aligned}
B(z^{-1}) &= b_0 + b_1 z^{-1} + \ldots + b_q z^{-q} \\
C(z^{-1}) &= 1 + c_1 z^{-1} + \ldots + c_r z^{-r}
\end{aligned}
\tag{3.59}
$$

and

$$
u_n = y_{n+d} = w_{n+d} = 0 \ , \quad \forall t \leq 0.
$$

Set

$$
\theta = (a_1, \ldots, a_p, b_0, \ldots, b_q, c_1, \ldots, c_r, \alpha_{10}, \ldots, \alpha_{1m}, \alpha_{20}, \ldots, \alpha_{lm})^T
$$

and

$$
\begin{aligned}
x_n^0 \ = \ & (y_{n-1}, \ldots, y_{n-p}, u_{n-d}, \ldots, u_{n-q-d}, w_{n-1}, \\
& \ldots, w_{n-r}, y_{n-1}u_{n-d}, \\
& \ldots, y_{n-1}u_{n-m-d}, y_{n-2}u_{n-d}, \ldots, y_{n-l}u_{n-m-d} \ )^T .
\end{aligned}
$$

System (3.58) can then be written as

$$
y_n = \theta^T x_n^0 + w_n \tag{3.60}
$$

2. The extended least squares algorithm

The Extended Least Squares Method is applied to estimate the unknown parameter θ , and θ_n represents the estimate of θ at time n. The Recursive Algorithm is the same as that presented by Chen *et al.* (1996):

$$
K_n = R_{n-1}x_n / \left(1 + x_n^T R_{n-1} x_n\right) \tag{3.61}
$$

$$
R_n = R_{n-1} - K_n x_n^T R_{n-1} \tag{3.62}
$$

$$
\theta_n = \theta_{n-1} + K_n \left(y_n - x_n^T \theta_{n-1}\right) \tag{3.63}
$$

$$
e_n = y_n - x_n^T \theta_n \tag{3.64}
$$

where x_n is constructed by using e_{n-i} instead of w_{n-i} in x_n^0, $i = 1, \ldots, r$, and $e_n = 0$ when $n \leq 0$. The dimensions of the parameter vector θ and regression vector x_n^0 are h, and

$h = p + q + 1 + l(m+1) + r$. Let $R_0 = hI_h$, $\theta_0 = 0$. We may choose R_0 to be any h th order positive definite matrix and θ_0 to be any h-dimensional vector.

In this thesis, the norm of the vector x is defined as $\|x\| = \sqrt{x^T x}$. Set

$$r_0 = Tr\ R_0^{-1} \quad , \quad r_t = r_0 + \sum_{i=1}^{n} \|x_i\|^2 \tag{3.65}$$

then $r_n = r_{n-1} + \|x_n\|^2$ $(n = 1, 2, \ldots)$. The set F_n is the σ-algebra set generated by $\{w_s,\ s \leq n\}$, i.e. $F_n = \sigma\{w_s,\ s \leq n\}$; and $\{w_n\}$ is assumed to be a martingale difference sequence. For all $n = 2, 3, \ldots$ the following properties hold:

$$E(w_n/F_{n-1}) = 0 \quad \text{a.s.} \quad \forall t \geq 1 \ , \quad F_0 = \{\phi, \Omega\} \tag{3.66}$$

$$E(w_n^2/F_{n-1}) \leq \xi_0 r_{t-1}^{\epsilon} \quad \text{a.s.} \quad \forall t \geq 1 \ , \quad 0 < \xi_0 < \infty \ , \quad 0 \leq \epsilon < 1. \tag{3.67}$$

Here $E(w_n^2/F_{n-1})$ may be unbounded with respect to ω and n. Additionally, the input u_n is F_n-measurable and

$$E(u_n^2) \leq \infty \ , \qquad n = 0, 1, 2, \ldots \tag{3.68}$$

If u_n is a deterministic signal, then $E(|u_n|^2) = |u_n|^2 < \infty$. Let $b_n = e_n - w_n$ and $\tilde{\theta}_n = \theta - \theta_n$.

3. Convergence Analysis

Here, we study the strong consistency and the convergence rate of the Extended Least Squares Identification for systems (3.58).

We assume that $\lambda_{\max}^n$ and $\lambda_{\min}^n$ are the maximum and minimum eigenvalues of the matrix R_n^{-1}, then

$$(\lambda_{\min}^n)^h \leq \det R_n^{-1} \leq (\lambda_{\max}^n)^h \tag{3.69}$$

From (3.34), we have

$$\lambda_{\min}^n \leq \lambda_{\max}^n \leq r_n \leq h\lambda_{\max}^n \tag{3.70}$$

As for the linear time-invariant case, in order to get strong consistency, we need

$$r_n \to \infty \quad \text{as} \quad n \to \infty \quad \text{a.s.} \tag{3.71}$$

to guarantee the strong convergence of θ_n, for all values of θ_0 . Otherwise even if the limit of θ_n exists, it may depend on the initial value of θ_0. For a sufficiently large positive number x, we set

$$\log k(x) = \underbrace{\log\ \log \ldots\ldots \log(x)}_{k \text{ times}}$$

and for $\delta > 1$

$$L_k^{\delta}(x) = \log x \log 2x \ldots \log (k-1)(x)(\log k(x))^{\delta}.$$

Theorem 3.13

Suppose that the noise and input of the system (3.58) satisfy the conditions (3.66)-(3.67), and the recursive algorithm (3.61)-(3.64) is used to estimate the parameter vector. If the following conditions are met

(T2a) $C^{-1}(z^{-1}) - 1/2$ is strictly positive real and the zeros of $C(z)$ are all outside the unit circle

(T2b) r_n meets the condition (3.71) and there exists a natural number k and constant $c > \delta > 1$ such that

$$\lim_{n\to\infty} r_n^c L_k^c(r_n)/\lambda_{\min}^n = 0 \quad \text{a.s.} \ , \tag{3.72}$$

then

$$\lim_{n\to\infty} \theta_n = \theta \quad \text{a.s.} \tag{3.73}$$

$$\|\tilde{\theta}_n\| = \mathcal{O}\left(\sqrt{r_n^c L_k^c(r_n)/\lambda_{\min}^n}\right) \quad \text{a.s.} \tag{3.74}$$

and

$$\lim_{n\to\infty} \frac{1}{r_n} \sum_{i=1}^{n} b_i^2 = 0 \quad \text{a.s.} \tag{3.75}$$

The proof of Theorem 3.13 see Chen *et al.* (1996).

Corollary 3.6.

Under the assumptions of Theorem 3.13, if we alter (3.72) to the stricter inequality

$$\lambda_{\max}^n/\lambda_{\min}^n \le \xi_1 \ , \quad \text{a.s.} \quad \forall n \ge 1 \ , \quad 1 \le \xi_1 < \infty \ , \tag{3.76}$$

then (3.73) and (3.75) still hold, and relation (3.74) can be improved to become

$$\|\tilde{\theta}_n\| = \mathcal{O}\left(\sqrt{r_n^{c-1} L_k^c(r_n)}\right) \quad \text{a.s.} \tag{3.77}$$

Corollary 3.7.

Under the conditions of Theorem 3.13, suppose the inequality (3.67) is changed to

$$E(w_n^2/F_{n-1}) \leq \xi_2 \quad \text{a.s.} \quad \forall n \geq 1 \ , \quad 0 \leq \xi_2 < \infty. \tag{3.78}$$

Since (3.78) is stricter than (3.67), we can get a faster convergence rate; (3.73) and (3.75) still hold and

$$\|\tilde{\theta}_n\| = \mathcal{O}\left(\sqrt{L_k^c(r_n)/\lambda_{\min}^n}\right) \quad \text{a.s.} \tag{3.79}$$

3.3.6 Extended Least Squares Identification and Its Convergence Analysis of MISO Model

1. Statement of the problem

For s different input time series, a MISO system can be written mathematically in the form:

$$\begin{aligned} y_n + A_1 y_{n-1} + ... + A_p y_{n-p} \ = \ & B_{1_1} u_{1_{n-1}} + ... + B_{1_{n-q_1}} u_{1_{n-q_1}} \\ & + B_{2_1} u_{2_{n-1}} + ... + B_{2_{n-q_2}} u_{2_{n-q_2}} \\ & ..... \\ & + B_{s_1} u_{s_{n-1}} + ... + B_{s_{n-q_s}} u_{s_{n-q_s}} \\ & + w_n + C_1 w_{n-1} + ... + c_r w_{n-r} \end{aligned} \tag{3.80}$$

y_n and $u_{i_n}(n = 1, 2, 3, ...)$ are the output and i-th input of the system respectively. p is the order of the system; $\{w_n\}, (n = 1, 2, 3, ...)$ is a noise series and the restriction on it is the same as the equations (3.66)-(3.67) and A_i, B_{j_k} and $C_l (i = 1, ..., p; j = 1, ...s; k = 1, ..., q_j; (q_j \leq p); l = 1, ...r.)$ are unknown parameters to be estimated.

Let z^{-1} is a unit delay operator and

$$A(z^{-1}) = 1 + A_1 z^{-1} + ... + A_p z^{-p} \tag{3.81}$$

$$B_i(z^{-1}) = B_{i_1} z^{-1} + ... + B_{i_{q_i}} z^{-q_i}, \quad i = 1, 2, \ldots, s \tag{3.82}$$

$$C(z^{-1}) = 1 + C_1 z^{-1} + ... + C_r z^{-r} \tag{3.83}$$

The equation (3.80) can be written as follows:

$$A(z^{-1})y_n = \sum_{i=1}^{s} B_i(z^{-1})u_{i_n} + \epsilon_n \tag{3.84}$$

where $\epsilon_n = C(z^{-1})w_n$ is the system noise and set

$$\theta^T = [-A_1, ..., -A_p, B_{1_1}, ..., B_{1_{q_1}}, B_{2_1}, ..., B_{s_{q_s}}, C_1, ..., C_r] \tag{3.85}$$

$$x_n^T = [y_{n-1}, ..., y_{n-p}, u_{1_n}, ..., u_{1_{n-q_1}}, u_{2_n}, ..., u_{s_{n-q_s}}, e_{n-1}, ..., e_{n-r}] \tag{3.86}$$

$$e_n = y_n - \theta_n^T x_n \tag{3.87}$$

$$d = p + r + \sum_{i=1}^{s} q_i \tag{3.88}$$

here θ is the true parameter matrix, x_n is the regression vector consisted of the information of input, output and the estimation of system noise, e_n is the estimation of w_n and θ_n is the estimate of θ at time n . It is easy to see that (3.80) or (3.84) also can be written as

$$y_n = \theta^T x_n + C(z^{-1})w_n + e_n - C(z^{-1})e_n \tag{3.89}$$

The first term of the right side of equation (3.89) can be considered as the estimation of y_n (since we do not know the true θ, y_n is estimated by $\theta_n^T x_n$) and the remaining terms on the right hand side can be considered as a kind of filter of system noise. In order to identify the system parameter vector θ, we make use of the recursive algorithms (3.61)–(3.63). d is given by (3.88) and the regression vector x_n is determined by (3.86).

Remark 3.1. In order to prove the strong consistency of the (RLSM) easily, we set $R_0 = dI_d$. In fact, R_0 may be any $d * d$ positive definite matrix.

Theorem 3.14

Suppose that the noise and input of the system (3.80) satisfy the conditions (3.66)–(3.67), and the recursive algorithm (3.61)–(3.64) is used to estimate the parameter vector. If the following conditions are met.

(T3a) $C^{-1}(z^{-1}) - \frac{1}{2}I_m$ is strictly positive real and the zeros of $C(z^{-1})$ are all outside the unit circle

(T3b) r_n meets the condition (3.71) and there exists a natural number k and constant $c > \delta > 1$ such that

$$\lim_{n\to\infty} r_n^c L_k^c(r_n)/\lambda_{\min}^n = 0, \tag{3.90}$$

then

$$\lim_{n\to\infty} \theta_n = \theta \tag{3.91}$$

$$\|\tilde{\theta}_n\| = \mathcal{O}\left(\sqrt{r_n^c L_k^c(r_n)/\lambda_{\min}^n}\right) \tag{3.92}$$

and

$$\lim_{n\to\infty} \frac{1}{r_n} \sum_{i=1}^{n} \|b_i\|^2 = 0. \tag{3.93}$$

Some lemmas and their proofs are given before we prove the Theorem 3.14.

Lemma 3.2.
For the conditions (3.66) and (3.67), the Identification Algorithm (3.61)-(3.64) of system (3.80) has the following properties, for $n = 1, 2, 3, \ldots$

$$r_n = \mathrm{Tr}\ (R_n^{-1}) \tag{3.94}$$

$$r_n < \infty\ ,\quad E(\|y_n\|^2) < \infty\ ,\quad E(\|\theta_n\|^2) < \infty\ ,\quad E(\|e_n\|^2)) < \infty \quad \text{a.s.} \tag{3.95}$$

$$E(\|\tilde{\theta}_n\|^2) < \infty\ ,\quad E(\|b_n\|^2) < \infty \quad \text{a.s.} \tag{3.96}$$

$$\tilde{\theta}_n = \tilde{\theta}_{n-1} - R_{n-1}x_n(b_n + w_n) \tag{3.97}$$

$$x_n^T R_n x_n = \left(\det R_n^{-1} - \det R_{n-1}^{-1}\right) / \det R_n^{-1} \tag{3.98}$$

$$C(z^{-1})b_n = \tilde{\theta}_n^T x_n \tag{3.99}$$

Proof: The steps are outlined below.

(i) Using induction, from (3.61)-(3.64) and the Inverse Matrix Lemma, (3.94) can easily be shown to hold.

(ii) Since the initial conditions of y_n and e_n are finite and also is the $\|x_0\| = 0$, $\|\theta\| < \infty$, $\|\theta_0\| < \infty$ from the system (3.80) and RLSM (3.61)-(3.63), we can obtain (3.95) for all $n = 1, 2, 3, \ldots$.

(iii) The first inequality of (3.96) obviously holds, and from (3.95) we get $0 \leq \xi_0 r_n < \infty$ a.s. So we have

$$E(\|w_n\|^2) = E(E(\|w_n\|^2/F_{n-1})) \leq E(\xi_0 r_{n-1}^c) < \infty \ , \quad \forall n \geq 1 \tag{3.100}$$

and (3.96) holds.

(iv) From (3.61)-(3.64), we deduce that $\theta_n = \theta_{n-1} + R_{n-1}x_n e_n$, and from $\theta_n = \theta - \tilde{\theta}_n$ and $e_n = b_n + w_n$, (3.97) is obtained.

(v) Equation (3.98) can be deduced by considering the determinant of the block matrix

$$\begin{pmatrix} 1 & x_n^T \\ x_n & R_n^{-1} \end{pmatrix}$$

(vi) From (3.87), (3.89) and $b_n = e_n - w_n$, (3.99) is obtained.

Lemma 3.3.

Assume that Lemma 3.2 and condition (3.71) hold. Then there exists a natural number $N(k)$ such that for any natural number k and d, the dimension of the parameter vector,

$$d \log k(r_n) \geq \log k(\det R_n^{-1}) \qquad \forall n \geq N(k) \tag{3.101}$$

Lemma 3.4.

Under the conditions of Lemma 3.3., for any natural number k and some positive number $\delta > 1$, there exists $N(k)$ such that

$$\sum_{i=N(k)}^{\infty} x_i^T R_i x_i / L_k^{\delta}(r_i) < \infty \tag{3.102}$$

Lemma 3.5.

Under the conditions of Lemma 3.3, for any natural number k and some natural number $N(k)$ and positive number $\delta > 1$, let

$$V_n = Tr(\tilde{\theta}_n^T R_n^{-1} \tilde{\theta}_n)/r_n^c L_k^{\delta}(r_n) \ , \quad \forall n \geq N(k) \tag{3.103}$$

where ϵ is given by (3.67). Then

$$E(V_n/F_{n-1}) \leq V_{n-1} + 2\xi_0 x_n^T R_n x_n / L_k^\delta(r_n) + E(x_n^T\tilde{\theta}_n(x_n^T\tilde{\theta}_n - 2b_n)/F_{n-1})/r_n^c L_k^\delta(r_n),\ \forall n \geq N(k) \quad (3.104)$$

The proofs of Lemmas 3.2.–3.5. are similar to those in Chen, Zinober and Ruan (1996) and Chen and Ruan (1987). Before proving Theorem 3.14 we make a few comments.

Remark 3.2. If in system (3.80), $C(z^{-1}) = I_m$, i.e. $q = 0$, then the conditions (3.66), (3.67) and (T3a) are met. Whether the conclusion holds depends on the conditions (T3b). The result of Theorem 3.14 can be extended to the case of the system with white noise.

Remark 3.3. If $\lim_{N\to\infty} \sup_{n\leq N}(r_n/N) < \infty$, then (3.93) is equivalent to

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \|b_i\|^2 = 0.$$

Proof:

From (T3a), there are two positive constants k_1 and k_2 such that

$$S_n = \sum_{i=N(k)}^{n} x_i^T\tilde{\theta}_i \left(C^{-1}(z^{-1}) - (\frac{1+k_1}{2})I_m\right) x_i^T\tilde{\theta}_i + k_2 \geq 0\ ,\quad \forall n \geq N(k) \quad (3.105)$$

$$S_n = \sum_{i=N(k)}^{n} x_i^T\tilde{\theta}_i b_i - \left(\frac{1+k_1}{2}\right)\left(\sum_{i=N(k)}^{n} \|x_i^T\tilde{\theta}_i\|^2\right) + k_2 \geq 0\ ,\quad \forall n \geq N(k) \quad (3.106)$$

Let

$$\xi_n = V_n + 2S_n/r_n^c L_k^\delta(r_n)\ ,\quad \forall n \geq N(k) \quad (3.107)$$

then ξ_n is a non-negative F_n-measurable stochastic sequence and from (3.104), (3.106) and (3.107), we get

$$E(\xi_{n+1}/F_n) \leq \xi_n + \eta_n - \zeta_n$$

where

$$\eta_n = 2\xi_0 x_{n+1}^T R_{n+1} x_{n+1}/L_k^\delta(r_{n+1}) \quad (3.108)$$

and

$$\zeta_n = k_1 E\left(\|x_{n+1}^T\tilde{\theta}_{n+1}\|^2/F_n\right)/r_{n+1}^c L_k^\delta(r_{n+1}) \quad (3.109)$$

are non-negative stochastic sequences, and also are F_n-measurable, from Lemma 3.5. We know that

$$\sum_{n=N(k)}^{\infty} \eta_n < \infty.$$

By using the Almost Supermartingale Theorem (Robbins and Siegmund, 1971) we get

$$\lim_{n\to\infty} \xi_n = \xi < \infty \qquad \text{and} \qquad \sum_{n=N(k)}^{\infty} \zeta_n < \infty \tag{3.110}$$

Since $C^{-1}(z^{-1})$ is a stable transform function, from (3.105), (3.109) and (3.110) we have

$$\sum_{n=N(k)}^{\infty} E\left(\|b_{n+1}\|^2/F_n\right)/r_{n+1}^{\epsilon} L_k^{\delta}(r_{n+1}) < \infty \tag{3.111}$$

and

$$\sum_{n=N(k)}^{\infty} E\left(\|b_{n+1}\|^2/r_{n+1}^{\epsilon} L_k^{\delta}(r_{n+1})\right) < \infty. \tag{3.112}$$

Set

$$C_n = \sum_{i=N(k)}^{n} \left(\|b_{i+1}\|^2 - E(\|b_{i+1}\|^2/F_i)\right)/r_{i+1}^{\epsilon} L_k^{\delta}(r_{i+1}) \ , \quad \forall n \geq N(k) \tag{3.113}$$

then (C_n, F_n) is a martingale, and from (3.112) and (3.113), we get

$$\sup_{n\geq N(k)} E|C_n| \leq 2 \sum_{i=N(k)}^{\infty} E\left(\|b_{i+1}\|^2/r_{i+1}^{\epsilon} L_k^{\delta}(r_{i+1})\right) < \infty$$

According to the Martingale Convergence Theorem, we have

$$\lim_{n\to\infty} C_n = \sum_{i=N(k)}^{\infty} \left(\|b_{i+1}\|^2 - E(\|b_{i+1}\|^2/F_i)\right)/r_{i+1}^{\epsilon} L_k^{\delta}(r_{i+1}) < \infty \tag{3.114}$$

and from (3.111) and (3.114),

$$\sum_{n=N(k)}^{\infty} \|b_n\|^2/r_n^{\epsilon} L_k^{\delta}(r_n) < \infty. \tag{3.115}$$

From the Kronecker Lemma, and (3.115)

$$\lim_{n\to\infty} \frac{1}{r_n^{\epsilon} L_k^{\delta}(r_n)} \sum_{i=N(k)}^{n} \|b_i\|^2 = 0. \tag{3.116}$$

So (3.93) can be easily deduced since $r_n \geq r_n^c L_k^\delta(r_n)$. Similarly to the proof of (3.116), from (3.109) and (3.110) we get

$$\lim_{n\to\infty} \frac{1}{r_n^c L_k^\delta(r_n)} \sum_{i=N(k)}^{n} \|x_i^T \tilde{\theta}_i\|^2 = 0 \tag{3.117}$$

From the Schwarz Inequality and (3.106)

$$0 \leq S_n \leq k_2 + \sum_{i=N(k)}^{n} x_i^T \tilde{\theta}_i b_i \leq k_2 + \left((\sum_{i=N(k)}^{n} \|x_i^T \tilde{\theta}_i\|^2)(\sum_{i=N(k)}^{n} \|b_i\|^2) \right)^{1/2} \tag{3.118}$$

Then, making use of (3.117) and (3.118)

$$\lim_{n\to\infty} S_n / r_n^c L_k^\delta(r_n) = 0 \tag{3.119}$$

From (3.107), (3.110) and (3.109) we obtain

$$\lim_{n\to\infty} V_n = V < \infty \tag{3.120}$$

Since $Tr(\tilde{\theta}_n^T R_n^{-1} \tilde{\theta}_n) \geq \lambda_{\min}^n \|\tilde{\theta}_n\|^2$, from (3.103)

$$\|\tilde{\theta}_n\|^2 \leq V_n r_n^c L_k^\delta(r_n) / \lambda_{\min}^n \tag{3.121}$$

Then (3.91) and (3.92) can be deduced from (3.90), (3.120) and (3.121) and $c > \delta$.

Corollary 3.8.

Under the assumptions of Theorem 3.14, if we alter (3.90) to the stricter inequality

$$\lambda_{\max}^n / \lambda_{\min}^n \leq \xi_1 \ , \qquad \forall n \geq 1 \ , \quad 1 \leq \xi_1 < \infty \ , \tag{3.122}$$

then (3.91) and (3.93) still hold, and relation (3.92) can be improved to become

$$\|\tilde{\theta}_n\| = \mathcal{O}\left(\sqrt{r_n^{c-1} L_k^c(r_n)} \right). \tag{3.123}$$

Corollary 3.9.

Under the conditions of Theorem 3.14, suppose the inequality (3.67) is changed to

$$E(\|w_n\|^2 / F_{n-1}) \leq \xi_2 \qquad \forall n \geq 1 \ , \quad 0 \leq \xi_2 < \infty \tag{3.124}$$

Since (3.124) is stricter than (3.67), we can get a faster convergence rate; (3.91) and (3.93) still hold and

$$\|\tilde{\theta}_n\| = \mathcal{O}\left(\sqrt{L_k^c(r_n) / \lambda_{\min}^n} \right). \tag{3.125}$$

3.4 Conclusion

In this chapter, some theoretical background of system identification and recursive least squares method are introduced. The so called extended least squares method (ELSM) for the different systems (MIMO ARMAX, Bilinear System and MISO system) and its strong consistency and convergence rate are derived and presented. As we all know, the least squares estimation is an old statistical method, but it is still investigated and applied by many statisticians due to its simplicity and practical importance. The convergence results obtained within mathematical statistics cannot be directly applied to system identification, since the design matrix is no longer deterministic in constrast to the classical situation. However, the least squares method has been analyzed and applied to the system identification problem for a long time. Obviously, the problem of the convergence of the least squares estimates in the system identification must be treated specifically.

The consistency of least squares estimates for white noise is discussed in (Han-fu Chen, 1985), but in general, it is inconsistent for coloured noise as we mentioned earlier in this chapter. Here, we modify the design matrix by using a new regression vector (3.64) and leaving the algorithm (3.61–3.63) invariant. The strong consistency and convergence rate at the condition (3.90) which is weaker than the well known persistent excitation condition are obtained by constructing a new series of stochastic Lyapunov functions for the special coloured noise case. These results provide the theorectical guarantee for the time series modelling the current velocity and SPM concentration dynamic system presented later in Chapters 6 and 7.

Chapter 4

Model Validation and Order Determination

4.1 Introduction

In system identification both the determination of model structure and model validation are important aspects. An overparametrized model structure can lead to unnecessarily complicated computations for finding the parameter estimates and for using the estimated model. An underparametrized model may be very inaccurate. The purpose of this chapter is to present a basic method that can be used to find an appropriate model order.
In practice one often performs identification for an increasing set of model orders (or more generally, structure indices). Hence one must know when the model order is appropriate, i.e. when to stop. Needless to say, any real-life data set cannot be modelled exactly by a linear finite-order model. Nevertheless such models often give good approximations of the true dynamics. However, the methods for finding the 'correct' model order are based on the statistical assumption that the data come from a true system within the model class considered.
When searching for the 'correct' model order one can raise different questions, which are discussed as follows:

- Is a given model flexible enough?
- Is a given model too complex?

- Which model structure of two or more candidates should be chosen?

Note that such questions are also relevant to model reduction.
The question that is asked in this section can also be phrased as: "Is the model structure large enough to cover the true system?" There are basically two ways to approach this question:

- Use of plots and common sense.
- Use of statistical tests on the prediction errors (residuals)

$$\epsilon(\hat{\theta}_N) = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{N}\hat{\epsilon}_i^2 \tag{4.1}$$

where y_n is the real data, $\hat{y}_n$ is the model output and $\hat{\epsilon}_n^2$ is the prediction error at time n respectively.
We concentrate on the latter aspect here. There are several statistical tests on the prediction errors $\epsilon(\hat{\theta}_N)$. The prediction errors evaluated at the parameter estimate $\hat{\theta}_N$ are often called the residuals. To simplify the notation, the residuals will frequently be written as just $\epsilon(\hat{\theta})$. Several statistical tests on the prediction errors are given in section 4.2 and an approach to order determination, minimum eigenvalue ratio test (MERT) is given in section 4.3.

4.2 Some Useful Tests

The methods for model structure determination based on tests of the residuals in practice are tied to model structures and identification methods where the disturbances are explicitly modelled. We assume that $\hat{\epsilon}_n$ is a zero mean white noise. The tests are formulated for single output system. For multivariable systems the tests have to be generalized. This can be done in a fairly straightforward manner, but the discussion will be confined to the scalar case to avoid cumbersome notation. The statistical properties of the tests will be analyzed under the null hypothesis that the model assumptions actually satisfy. Thus all the distribution results presented below hold under a certain null hypothesis.

4.2.1 An Autocorrelation Test

This test is based on assumption that $\hat{\epsilon}_n$ is a zero mean white noise. If $\hat{\epsilon}_n$ is white noise then its covariance function is zero except at $\tau = 0$:

$$r_\epsilon(\tau) = 0 \qquad \tau \neq 0 \tag{4.2}$$

First construct estimates of the covariance function as

$$\hat{r}_\epsilon(\tau) = \frac{1}{N}\sum_{i=1}^{N-\tau} \hat{\epsilon}_{i+\tau}\hat{\epsilon}_i \tag{4.3}$$

Under the assumption that $\hat{\epsilon}_n$ is a zero mean white noise, it can be deduced that

$$\hat{r}_\epsilon(\tau) \longrightarrow 0, \tau \neq 0; \quad \hat{r}_\epsilon(0) \longrightarrow \lambda^2 = E\epsilon_n^2, \qquad N \longrightarrow \infty. \tag{4.4}$$

To get a normalized test quantity, consider

$$x_\tau = \frac{\hat{r}_\epsilon(\tau)}{\hat{r}_\epsilon(0)} \tag{4.5}$$

According to (4.4) one can expect x_τ to be small for $\tau \neq 0$ and N to be large provided $\hat{\epsilon}_N$ is white noise. However, what does 'small' mean? To answer that question a more detailed analysis is necessary. Define

$$r = \frac{1}{N}\sum_{i=1}^{N} \begin{pmatrix} \hat{\epsilon}_{i-1} \\ \vdots \\ \hat{\epsilon}_{i-m} \end{pmatrix} \hat{\epsilon}_i = \begin{pmatrix} \hat{r}_\epsilon(1) \\ \vdots \\ \hat{r}_\epsilon(m) \end{pmatrix} \tag{4.6}$$

where, for convenience, the inferior limit of the sums was set to 1 (for large N this will have a negligible effect). From Söderström *et al.* (1989), r is asymptotically Gaussian distributed.

$$\sqrt{N}r \xrightarrow{dist} \mathcal{N}(0, P) \tag{4.7}$$

where the covariance matrix is

$$P = \lim_{N\to\infty} Err^T \tag{4.8}$$

The (i,j) element of P, $(i,j = 1,...,m)$ can be evaluated as

$$P_{i,j} = \lambda^4 \delta_{i,j}$$

Hence

$$P = \lambda^4 I$$

The result (4.7) implies that

$$Nr^T P^{-1} r = Nr^T r/r_\epsilon^2(0) \xrightarrow{dist} \chi^2(m)$$

Hence

$$\frac{N}{\hat{r}_\epsilon^2(0)} \sum_{i=1}^{m} \hat{r}_\epsilon^2(i) = Nr^T r/\hat{r}_\epsilon^2(0) \xrightarrow{dist} \chi^2(m)$$

It should be stressed once more that the distribution of the statistics presented above holds under the null hypothesis H_0 (asserting that $\hat{\epsilon}_n$ is white). The typical way of using the test statistics for model validation may be described as follows. Consider the test quantity $Nr^T r/\hat{r}_\epsilon^2(0)$. Let x denote a random variable which is χ^2 distributed with m degrees of freedom. Define $\chi_\alpha^2(m)$ by

$$\alpha = P(x > \chi_\alpha^2(m)) \tag{4.9}$$

for some given α which typically is chosen between 0.01 and 0.1. Then, if

$Nr^T r/\hat{r}_\epsilon^2(0) > \chi_\alpha^2(m)$ reject H_0 (and thus invalidate the model)

$Nr^T r/\hat{r}_\epsilon^2(0) \leq \chi_\alpha^2(m)$ accept H_0 (and thus validate the model).

Evidently the risk of rejecting H_0 when H_0 holds (which is called the first type of risk) is equal to α. The risk of accepting H_0 when it is not true depends on how much the properties of the tested model differ from H_0. The second type of risk cannot, in general, be determined for the statistics introduced previously, unless one restricts considerably the class of alternative hypotheses against which H_0 is tested. Thus, in applications the value of α (or equivalently, the value of the test threshold) should be chosen by considering only the first type of risk. When doing so it should, of course, be kept in mind that when α decreases the first type of risk also decreases, but the second type of risk increases. A frequently used value of α is $\alpha = 0.05$.

Remark 4.1

The number m in the above autocorrelation test could be chosen from 5 up to $N/4$. (see Söderstörm *et al.* (1989)).

4.2.2 The Parsimony Principle and the F-test

1. The Parsimony Principle

The parsimony principle is a useful rule when determining an appropriate model order. This principle says that out of two or more competing models which all explain the data well, the model with the smallest number of independent parameters should be chosen. Such a choice: 'Do not use extra parameters for describing a dynamic phenomenon if they are not needed'. The parsimony principal is discussed in Box and Jenkins (1976) and a theoretical justification is given in Söderstörm *et al.* (1989).

2. The F-test

The so called F-test can be used to compare two or more model structures. For such comparisons a discriminating criterion is needed. When the model structure is expanded so that more parameters are included in the parameter vector, the minimal value of loss function $\epsilon_N(\hat{\theta})$ naturally decreases since new degrees of freedom have been added to the optimization problem, or, in other words, the set over which optimization is done has been enlarged. The comparison of model structures can be interpreted as a test for significant decrease in the minimal values of the loss function associated with the (nested) model structures in question.

Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be two model structures, such that $\mathcal{U}_1 \subset \mathcal{U}_2$ ($\mathcal{U}_1$ is a subset of $\mathcal{U}_2$; for example, $\mathcal{U}_1$ corresponds to a lower-order model than $\mathcal{U}_2$). In such a case they are called hierarchical model structures. Further let $V_N^i = \epsilon_N^i(\hat{\theta})$ in the structure $\mathcal{U}_i$ ($i = 1, 2$) and let $\mathcal{U}_i$ have p_i parameters. We take

$$x = N\frac{V_N^1 - V_N^2}{V_N^2} \tag{4.10}$$

as a test quantity for comparing the model structures $\mathcal{U}_1$ and $\mathcal{U}_2$. If x is 'large' then we conclude that the decrease in the loss function from V_N^1 to V_N^2 is significant and hence the model structure $\mathcal{U}_2$ is significantly better than $\mathcal{U}_1$. On the other hand, when x is 'small', the conclusion is that $\mathcal{U}_1$ and $\mathcal{U}_2$ are almost equivalent and according to the parsimony principle the smaller model structure $\mathcal{U}_1$ should be chosen as the more appropriate one.
The discussion above leads to a qualitative procedure for discriminating between $\mathcal{U}_1$ and $\mathcal{U}_2$. To get a quantitative test procedure it is necessary to be more exact about what is meant by saying that x is 'large' or 'small'. This is done in the following.
First consider the case when $\mathcal{U}_1$ is not large enough to include the true system. Then the decrease $V_N^1 - V_N^2$ in the criterion function will be $O(1)$ (that is, it does not go to zero as $N \to \infty$) and therefore the test quantity x, (4.10), will be of magnitude N.
Next assume that $\mathcal{U}_1$ is large enough to include the true system, then it is possible to

prove that (see Söderstörm *et al.* 1989)

$$x = N\frac{V_N^1 - V_N^2}{V_N^2} \xrightarrow{dist} \chi^2(p_2 - p_1) \tag{4.11}$$

The result (4.11) can be used to conceive a simple test for model structure selection. At a significance level α (typical values in practice could range from 0.01 to 0.1) the smaller model structure $\mathcal{U}_1$ is selected over $\mathcal{U}_2$ if

$$x \leq \chi_\alpha^2(p_2 - p_1) \tag{4.12}$$

where $\chi_\alpha^2(p_2 - p_1)$ is defined by (4.11). Otherwise $\mathcal{U}_2$ is selected.

3. AIC and FPE Criteria

Another approach to model structure selection consists of using a criterion for assessment of the model structures under study. Such a criterion may for example be obtained by penalizing in some way the decrease of the loss function $\epsilon_N(\hat{\theta})$ with increasing model sets. The model structure giving the smallest value of this criterion is selected. Two forms of criterion are given in the following:

(i) AIC Criterion (Akaike's Information Criterion)

$$AIC = N log\ \epsilon_N(\hat{\theta}_N) + 2p \tag{4.13}$$

where p is the number of the parameters in the model.

(ii) The FPE Criterion (Final Prediction Error Criterion)

$$FPE = \epsilon_N(\hat{\theta}_N)\frac{1 + p/N}{1 - p/N} \tag{4.14}$$

It should be pointed out that model structure determination and model validation are very important step in system identification. For the determination of an appropriate model structure, it is recommended to use a combination of statistical tests.

4.3 An Approach to Order Determination

The specific and important problem of time series model order estimation has received considerable attention for many years and there are now a number of quite sophisticated

systematic procedures which appear to work quite well: (see e.g Box and Jenkins (1976); Åström and Eykhoff (1971); Akaike (1970, 1972, 1974); Young *et al.* (1980)). Of these, the final prediction error (FPE) and the information criterion (AIC) suggested by Akaike for purely stochastic processes are probably the most interesting in the present context. But unfortunately all these criteria may give more than one minimum, and may depend on assumptions constrained in the data and sometimes indicate too many parameters. Thus they should be used only as guides (Chatfield, 1980).

In practice, the time-series analyst often requires a simple yet robust procedure which is of a more general nature and emerges naturally during the time series analysis. Such a procedure has been suggested by Wellstead (1978) who develops an instrumental form of the determinant ratio test for model order based on the 'instrumental product-moment matrix' (IPM). He surveys previous related work on model order identification and shows that his approach is a natural development of previous order tests based on the product-moment matrix. It is however, less vulnerable to distortion if there is noise on the time-series data. Young *et al.* (1980) observed that a particular IPM occurs naturally in instrumental variable estimation. As a result, the procedure for generating the matrix suggested by Wellstead (1978) can be replaced by a more systematic but computationally more expensive procedure in which the instrumental variable sequence is obtained directly from the 'auxiliary model' used in the IV (instrumental variable) estimation. Alternatively, order estimation can be based on the inverse of the IPM, which occurs naturally in the IV algorithm and is directly related to the covariance matrix of the estimation errors. This interpretation provides an intuitively pleasing statistical explanation of the proposed order estimation procedure and it leads to the definition of a comprehesive model order identification procedure which is applicable to a wide variety of different models.

However, as it is well known, one of the fastest and simplest of the approximate order evaluation tools is based upon the product moment matrix of observed input/output data. The condition is induced by the linear constraint which defines the underlying system and hence provides a means of determining the system order. When estimating the order of unknown system prior to a parameter estimation, it is necessary to develop some useful methods to access roughly the order using a computationally simple and rapid algorithm. An improved estimate as well as verification of the underlying system order can then be obtained by combining some more sophisticated systematic procedures mentioned in the early part of this section.

The approximate order evaluation tool based upon the product moment matrix of observed data is originally due to Lee (1964), who pointed out that a simple rank condition exists for this matrix. The condition is induced by the linear constraint which defines the

underlying system and hence provides a means of determining the system order. Woodside (1971) exploited this rank condition to formulate a series of order determination algorithms. Subsequently other authors, notably Chow (1972); Tse and Wiener (1973) have applied this idea to autoregressive, moving average processes and the more challenging problems associated with structure determination of multivariable systems.
A major disadvantage of order tests based upon the product moment matrix is that the rank condition only applies to exactly observed data. In the presence of extraneous noise the condition for overparameterization is that the product moment matrix is 'almost singular'.

The aim of this subsection is to indicate an alternative simple and efficient test to ensure the rank condition has the meaning of mathematical expectation, and also to do this without significant additional computation or prior knowledge of the statistical properties of the system noises. The product moment matrix is then used by so-called minimum eigenvalue ratio test (MERT) in place of the normal determinant ratio test (DRT) suggested in Woodside (1971) to overcome both the expensive computational procedure and the effects of the system noise. Some examples and simulation results that are given in section 4 show that MERT is a simple and efficient test both in the linear system and in a class of nonlinear system. Also in the same section some comparisons between DRT and MERT are given.

4.3.1 Product Moment Matrix

The approach to model order estimation described here can be applied to most time-series models with both deterministic and stochastic input or 'exogenous' variables. To illustrate the method, let us first consider the discrete-time linear model for single input-single output (SISO) system which is the simpler form of ARMAX model in chapter 3 as follows:

$$A(z^{-1})y_n = B(z^{-1})u_n \tag{4.15}$$

where y_n and u_n are the observed system input and output sequence respectively and

$$v_n = u_n + r_n \qquad x_n = y_n + e_n \tag{4.16}$$

where v_n and x_n are input and output data; e_n and r_n are zero mean uncorrelated sequences with variances σ_e^2 and σ_r^2 respectively which are the source of all stochastic disturbances to the system; and $A(z^{-1}), B(z^{-1})$ are nth order polynomials in the backward

shift operator z^{-1} i.e.

$$A(z^{-1}) = 1 + a_1 z^{-1} + ... + a_n z^{-n}$$
$$B(z^{-1}) = b_1 z^{-1} + ... + b_n z^{-n} \tag{4.17}$$

Here all the polynomials are set equal to the order n of the characteristic polynomial A both for notational simplicity and to emphasise that the order of the characteristic polynomial $A(z^{-1})$ is the 'dynamic' order of the system (i.e. the order of the state space associated with the deterministic part of the system). In all that follows, we shall refer to n as the 'true' model order and denote the estimated model order as $\hat{n}$.

Our interest is to determine the integer k which characterises the order of the system in the presence of r_k and e_k. In the normal product moment formulation this is done by checking the rank of the $2(k+1) \times 2(k+1)$ matrix Γ_k defined by:

$$\Gamma_k(u,y) = \Omega_k^T(u,y)\Omega_k(u,y) \tag{4.18}$$

where

$$\Omega_k(u,y) = \begin{bmatrix} u_k & \dots & u_0 & y_k & \dots & y_0 \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ u_N & \dots & u_{N-k} & y_N & \dots & y_{N-k} \end{bmatrix} \tag{4.19}$$

and N is the number of observations.

If e_k and r_k are zero, the equation (4.19) gives an exact linear constraint on the data, the rank of $\Gamma_n(v,x)$ equals zero. Thus the system order is found by increasing the integer k until $\Gamma_n(v,x)$ becomes singular. In practice, however, the extraneous disturbance r_k and e_k will almost always be non-zero, such that the exact linear dependence in $\Omega_k(u,y)$ required for rank collapse cannot be set up. The practical consequence is that for $k > n$, the minimum eigenvalue or determinant of $\Gamma_k(u,y)$ falls to some lower bound determined by the statistical properties of the disturbances e_k and r_k. From (4.17)-(4.19),

$$\begin{aligned} E[\Gamma_k(u,y)] &= E(\Omega_k(v,x) + \Omega_k(r,e))^T(\Omega_k(v,x) + \Omega_k(r,e)) \\ &= E[\Gamma_k(v,x)] + R_k(r,e) \end{aligned} \tag{4.20}$$

where $R_k(r,e,v,x) = E[\Gamma_k(r,e) + \Omega_k(v,x)^T\Omega_k(r,e) + \Omega_k(r,e)^T\Omega_k(v,x)]$.

4.3.2 Minimum Eigenvalue Ratio Test

Theorem 4.1.

If **A** *is a non-negative symmetric matrix, the following conditions are equivalent:*

(I) $\mathbf{A}$ *is a singular matrix*
(II) det $\mathbf{A}=0$
(III) $\lambda_{min}(\mathbf{A})=0$
where det $\mathbf{A}$ *and* $\lambda_{min}(\mathbf{A})$ *denote the determinant and the minimum eigenvalue of matrix* $\mathbf{A}$ *respectively.*

Theorem 4.2.
If $\mathbf{A}$ *is a nth order non-negative symmetric matrix, there exists an orthogonal matrix* $\mathbf{T}$ *such that:*

$$\mathbf{T}^{\tau}\mathbf{A}\mathbf{T} = diag(\lambda_n, \ldots, \lambda_1) \tag{4.21}$$

where $\lambda_n \geq \lambda_{n-1} \geq \ldots \geq \lambda_1$ *are n eigenvalues of matrix* $\mathbf{A}$ *and* $\mathbf{T}^{\tau}$ *is the transpose matrix of* T.

Theorem 4.3.
If A and B are two symmetric matrices, then

$$\lambda_{min}(A) + \lambda_{min}(B) \leq \lambda_{min}(A+B) \leq \lambda_{min}(A) + \lambda_{max}(B)$$

(see Suda *et al.* (1973)).

Theorem 4.4
(i) If $\{r_k\}$ *and* $\{e_k\}$ *are white noise disturbances and independent of* $\{v_k\}$ *and* $\{x_k\}$ *then* $R_k(r,e,v,x) = diag(\underbrace{\sigma_r^2, \ldots, \sigma_r^2}_{k+1}, \underbrace{\sigma_e^2, \ldots, \sigma_e^2}_{k+1})$
(ii) If $\{r_k\}$ *and* $\{e_k\}$ *are white noise disturbances, independent of* $\{v_k\}$ *and* $\{x_k\}$ *and* $\sigma_e^2 = \sigma_r^2$, *then* $R_k(r,e,v,x) = diag(\sigma_r^2, \ldots, \sigma_r^2)$ *and*

$$\lambda_{min}(E\Gamma_k(u,y)) = \lambda_{min}(E\Gamma_k(v,x)) + \sigma_r^2. \tag{4.22}$$

(iii) If $\{r_k\}$ *and* $\{e_k\}$ *are white noise disturbances and* $\sigma_e^2 \neq \sigma_r^2$, *then*

$$\lambda_{min}(E\Gamma_k(u,y)) = \lambda_{min}(E\Gamma_k(v,x)) + R_k,$$
$$min(\sigma_e^2, \sigma_r^2) \leq R_k \leq max(\sigma_e^2, \sigma_r^2) \tag{4.23}$$

(iv) If $\{r_k\}, \{e_k\}$ *are coloured noise series and there exists a constant* C *such that:*

$$\lambda_{max}(R_k(r,e,x,v)) \leq C, \quad k = 1, 2, \ldots \tag{4.24}$$

then

$$\lambda_{min}(E\Gamma_k(u,y)) = \lambda_{min}(E\Gamma_k(v,x)) + R_k, \qquad R_k \leq C. \tag{4.25}$$

Proof:

(i) The proof of this follows from the assumption that e_k and r_k are uncorrelated white noise series and independent of $\{v_k\}$ and $\{x_k\}$.

(ii) Since $E\Gamma(u,y)$ is a non-negative symmetric matrix, from Theorem 4.2, there exists a orthogonal matrix $\mathbf{T}$, such that

$$\mathbf{T}^T E\Gamma_k(v,x)\mathbf{T} = diag(\lambda_1,\ldots,\lambda_{2k}) \tag{4.26}$$

where λ_i is the ith eigenvalue of the matrix $E\Gamma_k(v,x)$. So

$$\begin{aligned}\mathbf{T}^T E\Gamma_k(u,y)\mathbf{T} &= \mathbf{T}^T E\Gamma_k(v,x)\mathbf{T} + diag(\sigma_r^2,\ldots,\sigma_r^2).\\ &= diag(\lambda_1+\sigma_r^2,\ldots,\lambda_{2k}+\sigma_r^2)\end{aligned} \tag{4.27}$$

Therefore (4.22) holds.

$$\begin{aligned}&\frac{\lambda_{min}(E\Gamma_k(u,y)) - \lambda_{min}(E\Gamma_{k-1}(u,y))}{\lambda_{min}(E\Gamma_{k+1}(u,y)) - \lambda_{min}(E\Gamma_k(u,y))}\\ =\ &\frac{\lambda_{min}(E\Gamma_k(v,x)) - \lambda_{min}(E\Gamma_{k-1}(v,x))}{\lambda_{min}(E\Gamma_{k+1}(v,x)) - \lambda_{min}(E\Gamma_k(v,x))}\end{aligned} \tag{4.28}$$

Since $E\Gamma_k(v,x) = 0, (k \geq n)$ and $E\Gamma_k(v,x) \neq 0, (k < n)$, (14) become infinite.

(iii) According to Theorem 4.3, (4.23) holds.

If $k = n$, $\lambda_{min}E\Gamma_k(v,x) = 0$, then $\lambda_{min}E\Gamma_k(u,y)$ is limited its range being between σ_r^2 and σ_e^2. In practice,

$$\frac{\lambda_{min}E\Gamma_{k-1}(u,y) - R_{k-1}}{\lambda_{min}E\Gamma_k(u,y) - R_k} \tag{4.29}$$

should be very large. Since we assume that σ_r^2 and σ_e^2 are usually very small and the differences of $R_i, i = k-1, k, k+1$ are also very small,

$$\begin{aligned}&\frac{\lambda_{min}(E\Gamma_k(u,y)) - \lambda_{min}(E\Gamma_{k-1}(u,y))}{\lambda_{min}(E\Gamma_{k+1}(u,y)) - \lambda_{min}(E\Gamma_k(u,y))}\\ \sim\ &\frac{(\lambda_{min}E\Gamma_{k-1}(u,y) - R_{k-1}) - (\lambda_{min}E\Gamma_k(u,y) - R_k)}{(\lambda_{min}E\Gamma_k(u,y) - R_k) - (\lambda_{min}E\Gamma_{k+1}(u,y) - R_{k+1})}\\ =\ &\frac{\lambda_{min}(E\Gamma_k(v,x)) - \lambda_{min}(E\Gamma_{k-1}(v,x))}{\lambda_{min}(E\Gamma_{k+1}(v,x)) - \lambda_{min}(E\Gamma_k(v,x))}\end{aligned} \tag{4.30}$$

should be comparatively large since $E\Gamma_k(v,x) = 0, (k \geq n)$ and $E\Gamma_k(v,x) \neq 0, (k < n)$. It is noticed that (4.30) is equivalent to (4.28).

(iv) This is proved in a similar fashion to (iii). Theorem 4.4 is thus established.

From Theorems 4.1-4.4, we can easily show that the minimum eigenvalue of the product

moment matrix is very sensitive to its singularity. This is one main reason that we choose the minimum eigenvalue ratio test for estimating the order of the system. Also let us illustrate the reason that we choose the eigenvalue ratio rather than the determinant ratio test of Woodside (1971) by following simple examples: set a matrix

$$M_n = \begin{bmatrix} e^{n^2} & 0 \\ 0 & e^{-n^2} \end{bmatrix} \tag{4.31}$$

When n is sufficiently large, we can say M_n is 'nearly singular' due to its minimum eigenvalue e^{-n^2} being very very small. But unfortunately,

$$detM_n \equiv 1.$$

So the singularity information can not be found just from the deteminant in this case. Let

$$DR(k) = \frac{detM_k}{detM_{k+1}} \equiv 1$$

Meanwhile from M_n, we can easily see that the minimum eigenvalue of M_n is a much more sensible variable to reflect its singularity than the determinant of M_n. Let

$$\begin{aligned} MER(k) &= \frac{\lambda_{min}(M_{k-1}) - \lambda_{min}(M_k)}{\lambda_{min}(M_k) - \lambda_{min}(M_{k+1})} \\ &= \frac{e^{-(n-1)^2}(1 - e^{-(2n-1)})}{e^{-n^2}(1 - e^{-(2n+1)})} \\ &\geq \frac{e^{n^2}}{e^{(n-1)^2}} \geq e^{2n-1} \to \infty \end{aligned}$$

Some more discussion and examples will be given in the next subsection. In practice, the minimum eigenvalue of $\Gamma_k(u,y)$ will almost always be non-zero, it is expected that for $k > n$ the minimum eigenvalue of $\Gamma_k(u,y)$ falls to some lower bound range controlled by the noise covariance matrix. The minimum eigenvalue ratio test is used in place of the determinant ratio test for the nearly singularity of the product matrix. The so called minimum eigenvalue test is as follows:

$$MER(k) = \frac{\lambda_{min}(\Gamma_{k-1}(u,y)) - \lambda_{min}(\Gamma_k(u,y))}{\lambda_{min}(\Gamma_k(u,y)) - \lambda_{min}(\Gamma_{k+1}(u,y))}. \tag{4.32}$$

The quantity MER(k) forms a normalized order test quantity, such that when $k = n$, MER(k) increase rapidly. In practice, we increase k one by one and if MER(k) reaches its maximum, we can consider $k = \hat{n}$ as the estimate of n. From (4.28), we can see that MER($\hat{n}$) means comparatively large differences between $\Gamma_{\hat{n}-1}(u,y)$ and $\Gamma_{\hat{n}}(u,y)$ and small differences between $\Gamma_{\hat{n}}(u,y)$ and $\Gamma_{\hat{n}+1}(u,y)$ which also occurs in the behaviour of

$\lim_{k\to\infty}\lambda_{min}(\Gamma_k(u,y)) = 0$ or $\lambda_{min}(\Gamma_k(u,y))$ approaches to its lower bound range. Now we consider a more general case i.e. a kind of bilinear system introduced by Chen *et al.* (1996) and we only consider a simpler form of bilinear system as follows:

$$A(z^{-1})y_n = B(z^{-1})u_n + C(z^{-1})u_n y_n \tag{4.33}$$

where y_n and u_n are the observed system input and output sequences respectively and

$$v_n = u_n + r_n \qquad x_n = y_n + e_n \tag{4.34}$$

v_n and x_n are input and output data; e_n and r_n are zero mean uncorrelated sequences with variance σ_e^2 and σ_r^2 respectively are the sources of all stochastic disturbances to the system; and $A(z^{-1}), B(z^{-1})$ and $C(z^{-1})$ are nth order polynomials in the backward shift operator z^{-1} i.e.

$$\begin{aligned} A(z^{-1}) &= 1 + a_1 z^{-1} + ... + a_n z^{-n} \\ B(z^{-1}) &= b_1 z^{-1} + ... + b_n z^{-n} \\ C(z^{-1}) &= c_1 z^{-1} + ... + c_n z^{-n}. \end{aligned} \tag{4.35}$$

Set:

$$\Gamma_k(u,y) = \Omega_k^T(u,y)\Omega_k(u,y) \tag{4.36}$$

where

$$\Omega_k(u,y) = \begin{bmatrix} u_k & \dots & u_0 & y_k & \dots & y_0 & u_k y_k & \dots & u_0 y_0 \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ u_N & \dots & u_{N-k} & y_N & \dots & y_{N-k} & u_N y_N & \dots & u_{N-k} y_{N-k} \end{bmatrix} \tag{4.37}$$

and N is the number of observations. Set:

$$\Lambda_k(v,x,r,e) = \begin{bmatrix} 0 & \dots & 0 & 0 & \dots & 0 & v_k e_k + x_k r_k & \dots & v_0 e_0 + x_0 r_0 \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & v_N e_N + x_N r_N & \dots & v_{N-k} e_{N-k} + x_{N-k} r_{N-k} \end{bmatrix} \tag{4.38}$$

$$\begin{aligned} E\Gamma_k(u,y) &= E(\Omega_k(v,x) + \Omega_k(r,e) + \Lambda_k(v,x,r,e))^T(\Omega_k(v,x) + \Omega_k(r,e) \\ &\quad + \Lambda_k(v,x,r,e)) = E\Gamma_k(v,x) + \tilde{R}(v,x,r,e) \end{aligned} \tag{4.39}$$

where

$$\begin{aligned} \tilde{R}(v,x,r,e) &= E(\Omega_k(v,x))^T(\Omega_k(r,e) + \Lambda_k(v,x,r,e)) \\ &\quad + (\Omega_k(r,e) + \Lambda_k(v,x,r,e))^T\Omega_k(v,x)). \end{aligned} \tag{4.40}$$

Theorem 4.5.

(i) If $\{r_k\}$ *and* $\{e_k\}$ *are white noise disturbances and independent of* $\{v_k\}$ *and* $\{x_k\}$, *then* $R_k(r,e,v,x) = diag(\underbrace{\sigma_r^2,\ldots,\sigma_r^2}_{k+1},\underbrace{\sigma_e^2,\ldots,\sigma_e^2}_{k+1},\underbrace{\sigma_e^2\sigma_r^2,\ldots,\sigma_r^2\sigma_e^2}_{k+1})$.

(ii) If $\{r_k\},\{e_k\}$ *are coloured noise series and there exists a constant* C *such that:*

$$\lambda_{max}(\tilde{R}_k(r,e,x,v)) \leq \sigma^2, \quad k = 1,2,\ldots \tag{4.41}$$

then

$$\lambda_{min}(E\Gamma_k(u,y)) = \lambda_{min}(E\Gamma_k(v,x)) + \tilde{R}_k, \qquad \tilde{R}_k \leq C. \tag{4.42}$$

The proof of the Theorem 4.5 is similar to that of those in Theorem 4.4. It is natural that the MERT can also be used in the bilinear system.

4.3.3 Examples

Three examples are presented here to illustration how the MERT works for the linear and bilinear system with white noise and coloured noise respectively.
First we consider the example given by Wellstead (1978) whereby we can compare Woodside's DRT .

Example 4.1.
We consider the system as follows:

$$y_{n+1} - 1.5y_n + 0.7y_{n-1} = u_n + 0.5u_{n-1} + e_{n+1} \tag{4.43}$$

where e_k is white noise with variance σ^2 and is independent of u_k and in here for simplicity we assume $r_k \equiv 0$. Table 4.1 shows the variation in the determinant ratios (DR) and the minimum eigenvalue ratio (MER) of the product moment matrix for various orders, and based upon 500 data points and the u_k is a white noise series with $Eu_k = 0$ and $\sigma_u = 1$.

TABLE 4.1. The Comparison of order estimate test of DR and MER

	$\sigma = 0$	$\sigma = 0.01$	$\sigma = 0.05$	$\sigma = 0.1$	$\sigma = 0.2$
DR(1)	3.51297	3.51359e-08	3.50882e-08	3.49439e-08	3.50009e-08
DR(2)	44276.1	0.0114784	0.000388697	0.000122614	3.22383e-05
DR(3)	141572	0.0115572	0.000391872	0.000123378	3.29072e-05
DR(4)	217848	0.0117354	0.000400107	0.000126035	3.33166e-05
MER(1)	4.39868	4.39864	4.38944	4.39958	4.36695
MER(2)	1.07078e+12	270243	8204.94	2858.12	638.784
MER(3)	2.38109	2.076311	2.40435	2.30346	3.080796
MER(4)	1.92904	2.6573	1.9434	2.019417	2.25812

Figure 4.1.1 shows the variation in the determinant ratios (DR) of the product moment matrix for various orders. Note that as the variance of e_k is increased the change in the product moment determinant ratio is blurred.
Figure 4.1.2 shows the variation in MER method of the product moment matrix for various orders. Note that as the variance of e_k is increased the change in the MER version retains, to a certain extent, its discriminatory power.

Example 4.2.
We consider the system with the coloured noise as follows:

$$y_{n+1} - 1.5y_n + 0.7y_{n-1} = u_n + 0.5u_{n-1} + e_{n+1} - 0.5e_n \tag{4.44}$$

Table 4.2 shows the variation in the determinant ratios (DR) and the minimum eigenvalue ratio (MER) of the product moment matrix for various orders, and based upon 500 data points and the u_k is a white noise series with $Eu_k = 0$, and $\sigma_u = 1$.

TABLE 4.2. The Comparison of order estimate test of DR and MER

	$\sigma = 0$	$\sigma = 0.01$	$\sigma = 0.05$	$\sigma = 0.1$	$\sigma = 0.2$
DR(1)	2.1956e-05	2.19217e-05	2.19444e-05	2.14101e-05	2.10853e-05
DR(2)	6.70777e+07	0.239791	0.0101751	0.00249238	0.000546805
DR(3)	7.83125e+07	0.284344	0.0122996	0.00286875	0.000665293
DR(4)	1.04709e+08	0.303033	0.0129332	0.00293136	0.000697109
MER(1)	4.39868	4.4041	4.44528	4.60215	5.26181
MER(2)	2.17086e+13	112292	19590.5	789.234	132.171
MER(3)	0.0844448	0.128343	0.0416726	0.544485	0.256149
MER(4)	1.22852	5.14429	1.62926	3.047697	1.2321

Figures 4.2.1–4.2.2 show the similar results as Figures 4.1.1–4.1.2.

Example 4.3.

We consider a bilinear system similar to that of Chen *et al.* (1996) as follows:

$$\begin{aligned} y_{n+1} - 1.5y_n \quad &+ \quad 0.7y_{n-1} \\ &= \quad u_n + 0.5u_{n-1} - 0.2y_n u_n + 0.1y_{n-1}u_{n-1} + e_{n+1} + 0.2e_n \end{aligned} \tag{4.45}$$

Table 4.3 shows the variation in the determinant ratios (DR) and the minimum eigenvalue ratio (MER) of the product moment matrix for various orders, based once again upon 500 data points and the u_k is a white noise series with $Eu_k = 0$, and $\sigma_u = 1$.

TABLE 4.3. The Comparison of order estimate test of DR and MER

	$\sigma = 0$	$\sigma = 0.01$	$\sigma = 0.05$	$\sigma = 0.1$	$\sigma = 0.2$
DR(1)	1.4945e-09	1.49625e-09	1.52628e-09	1.48044e-09	1.45691e-09
DR(2)	4862.63	3.24106e-05	1.49464e-06	2.998e-07	7.22596e-08
DR(3)	7534.74	3.44451e-05	1.56957e-06	3.1102e-07	7.53756e-08
DR(4)	34410.4	3.49498e-05	1.61273e-06	3.18014e-07	7.78108e-08
MER(1)	6.52662	6.52863	6.553	6.71938	7.83037
MER(2)	4.45992e+12	6269178	307.595	60.3803	13.06294
MER(3)	0.280472	3.23657	3.083804	2.96004	2.55572
MER(4)	0.621848	2.27728	2.66296	2.42319	2.35324

Figures 4.3.1–4.3.2 show the similar results as Figures 4.1.1–4.1.2.
From Tables 4.1.–4.3. and Figures 4.1.1–4.3.2, we deduce the following:

- When there are no disturbances ($\sigma = 0$), both DRT and MERT work very well.
- When $\sigma \neq 0$, MERT works very well but not DRT, even if the variance of the noise is very small.
- The sensitivity of MERT decreases as σ increases, but considering the ratio of signal/noise $S = 10 \times \log \sigma_u/\sigma$, the high noise (S=16, i.e. $\sigma = 0.2$) case, the MERT looks reasonably good.
- The MERT works well both in bilinear and linear system with coloured and white noise disturbances.

From three examples presented here, we deduce that the MERT is a very practical method to estimate the order of linear and nonlinear system prior to parameter identification which can overcome both insensitivity of the system noise and computationally expensive procedures.

4.4 Conclusion

Some useful model validation and order determination tests and techniques are reviewed and the minimum eigenvalue ratio test of product moment matrix introduced in this chapter is a natural extension of the deterministic system. The purpose of presenting here is to provide a straightforward way to estimate approximately the order of unknown system as a complement of some popular methods reviewed in the first two sections of the chapter. The MERT method has the advantage of a crisp asymptotic rank condition which is not influenced by the observed input/output data disturbances and provide a simply and direct way to estimate the approxiamate order of a unknown system by using input/output data set. Moreover, this method can be extended to a first stage estimate on more complicated structures of the model as well as the order of a wide range system model. For the system with significant disturbance about the observed input/output data, more statistical reliability analysis should be carried out and it is suggested that

the method supposed to be carefullly used or combined with the comparatively mature methods reviewed in the chapter.

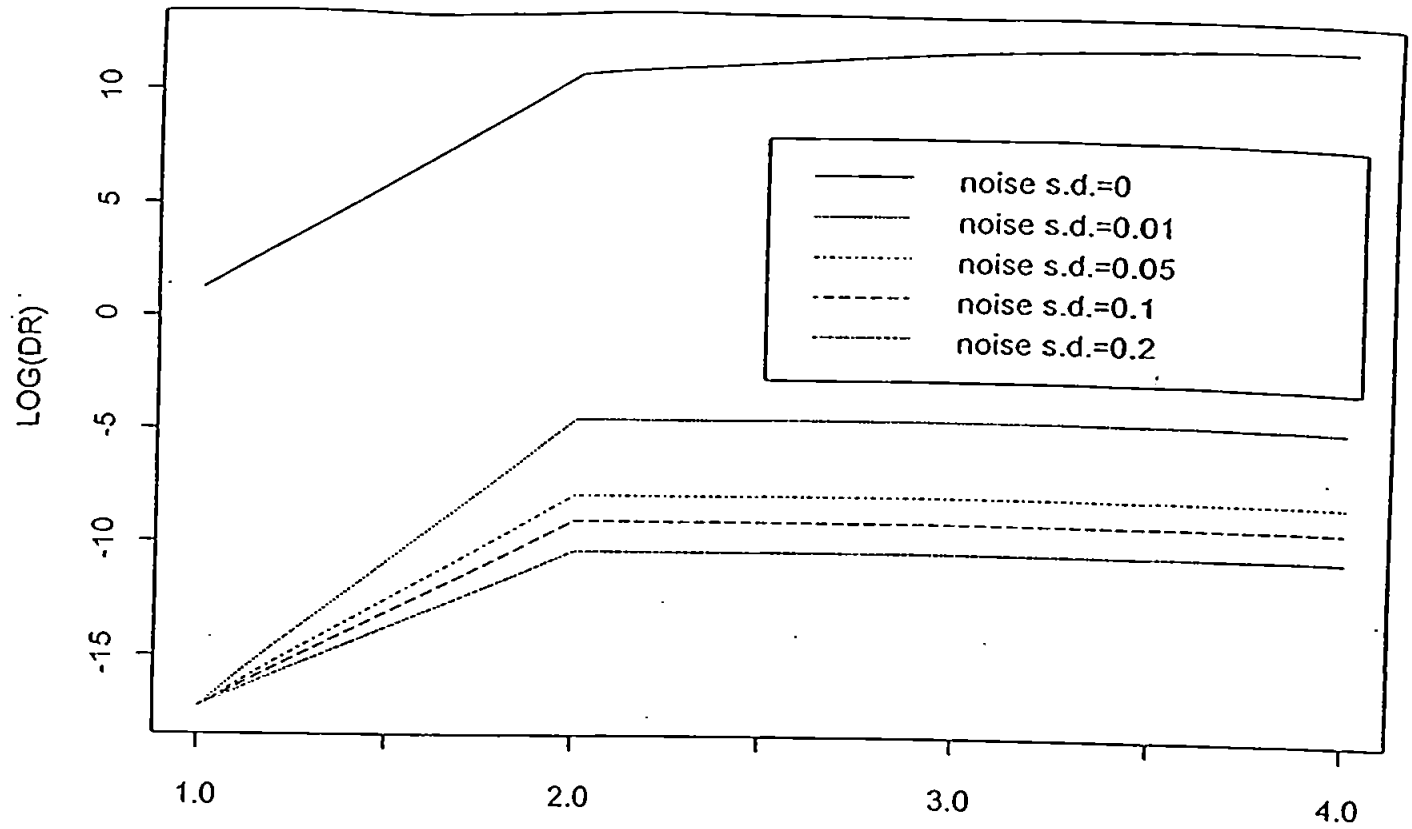


Figure 4.1.1. Variation of DR in Example 4.1

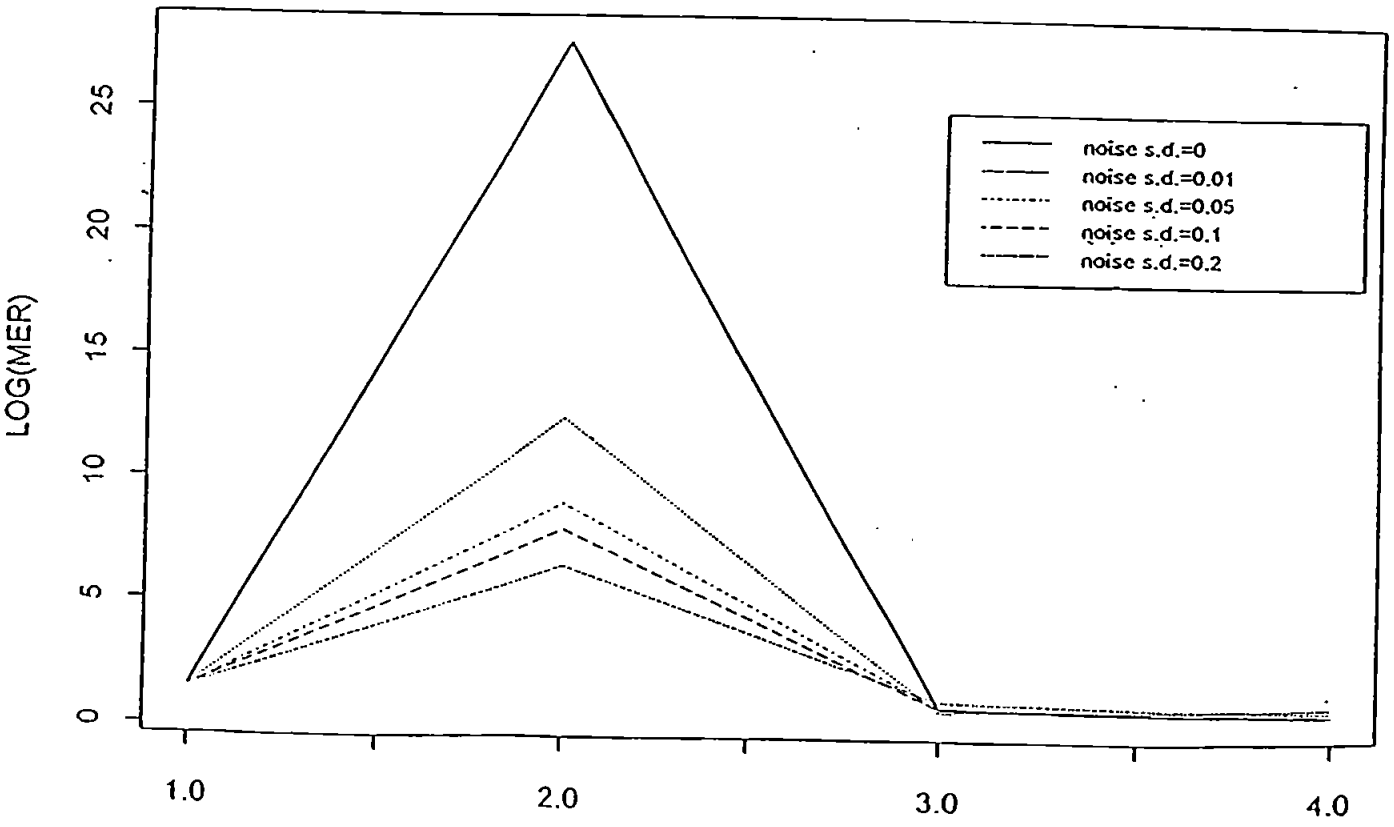


Figure 4.1.2. Variation of MER in Example 4.1

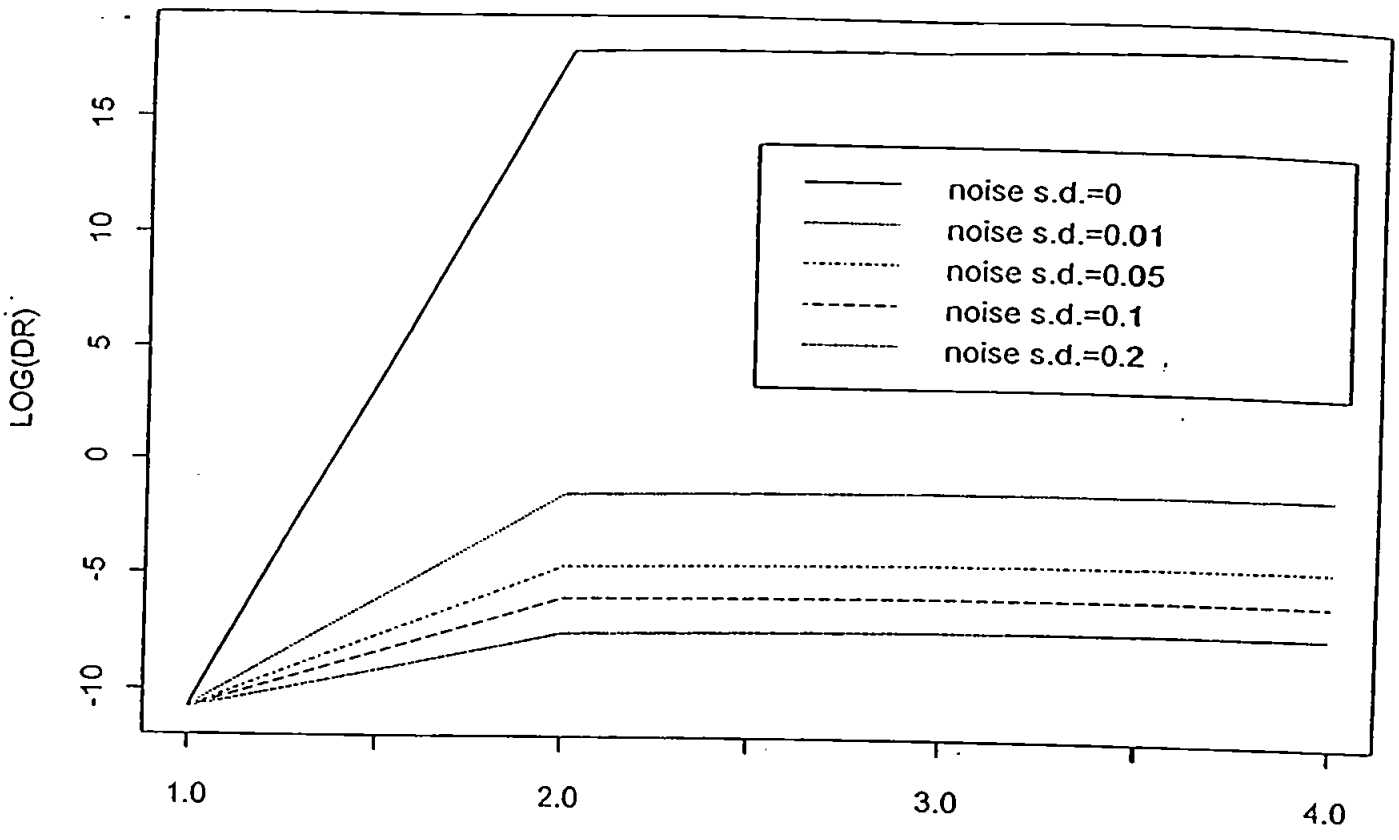


Figure 4.2.1. Variation of DR in Example 4.2

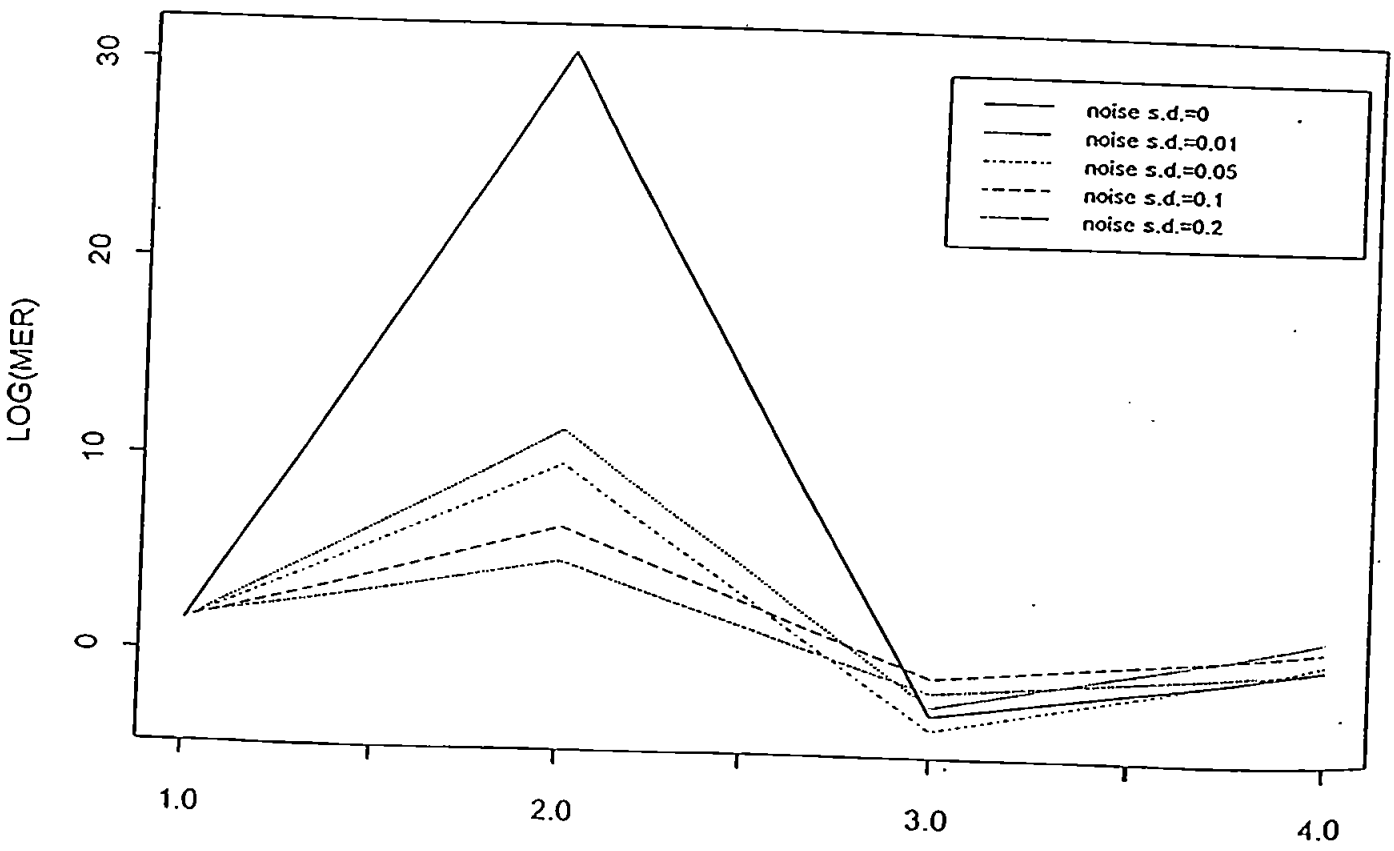


Figure 4.2.2. Variation of MER in Example 4.2

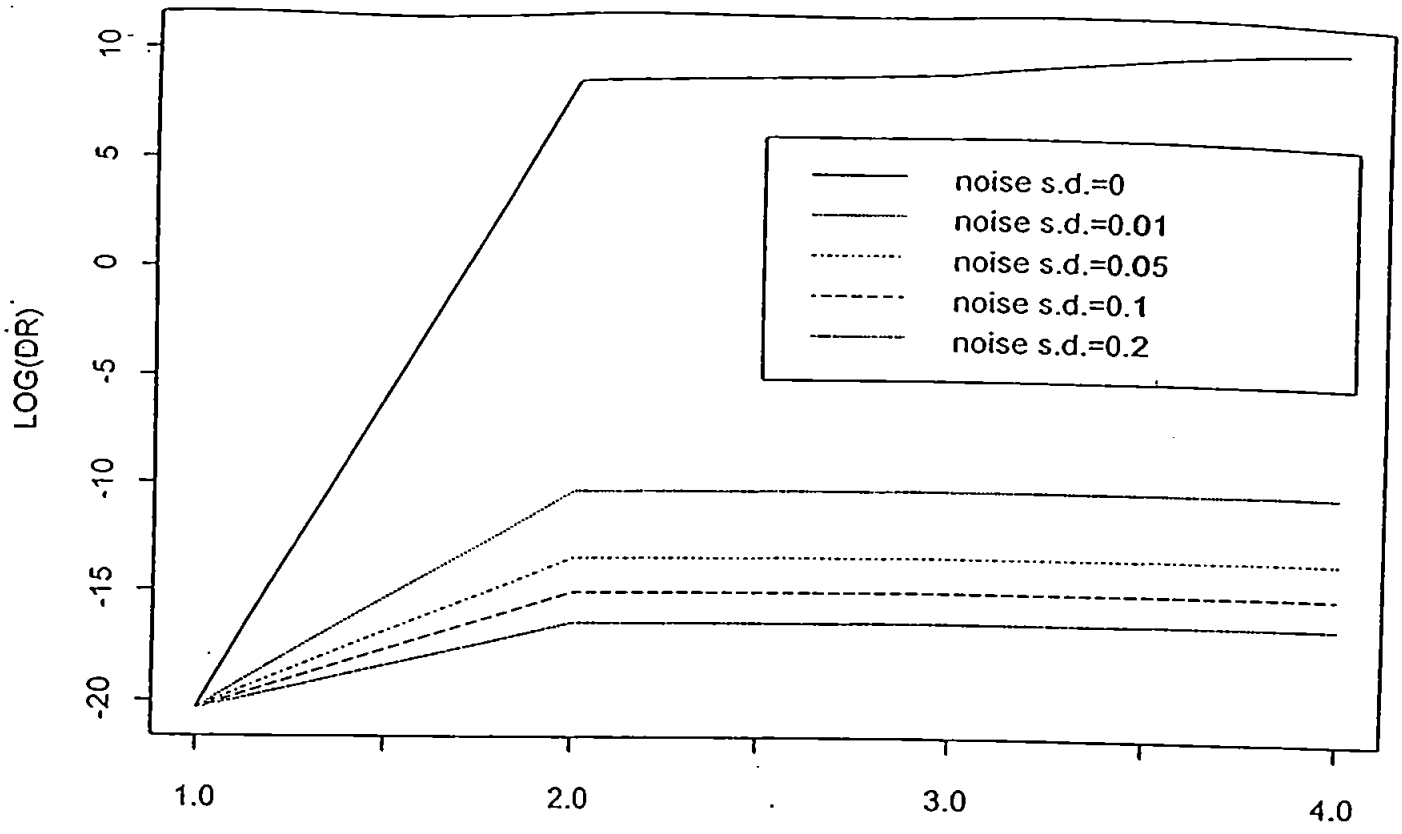


Figure 4.3.1. Variation of DR in Example 4.3

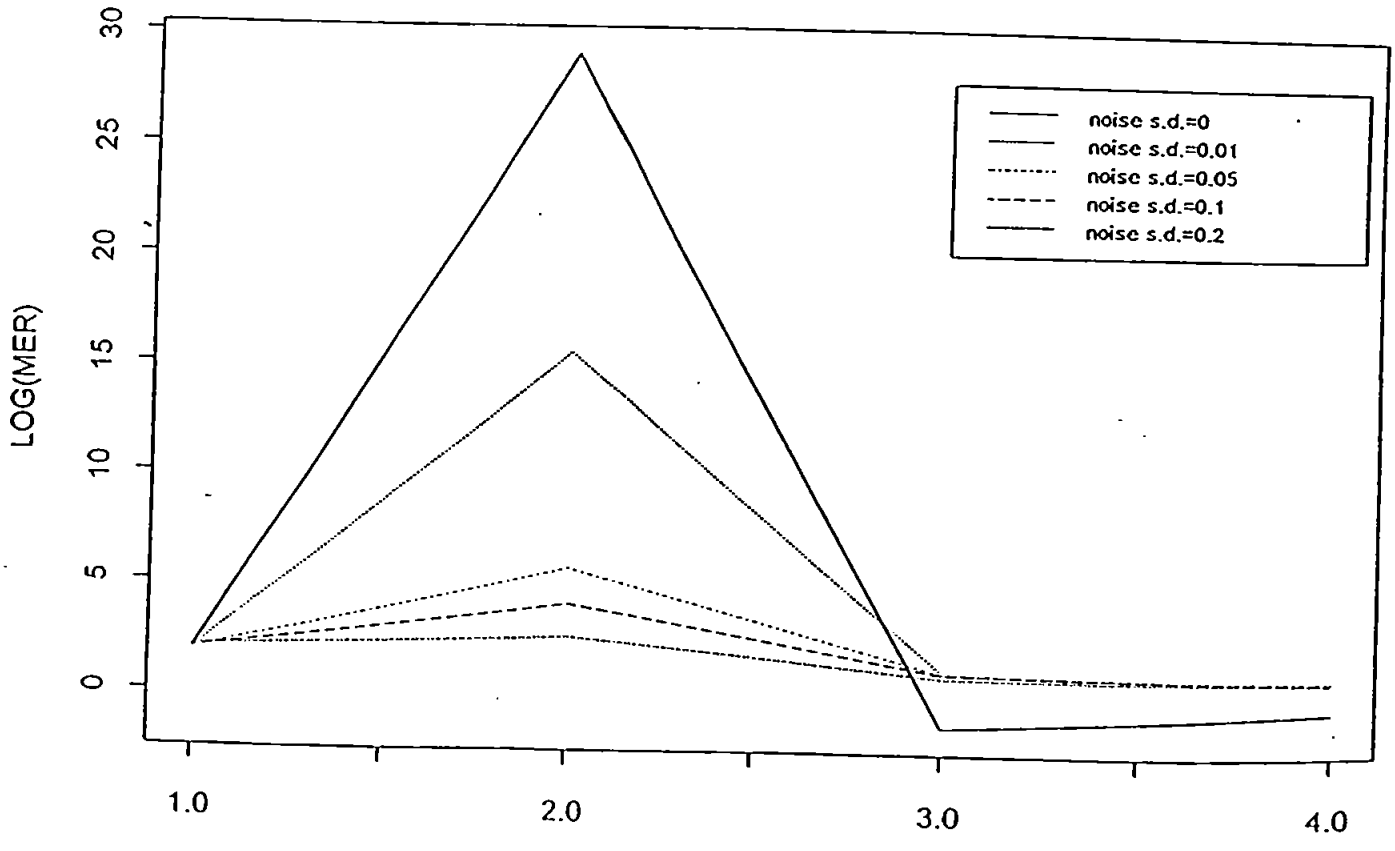


Figure 4.3.2. Variation of MER in Example 4.3

Chapter 5

Review of Cohesive Sediment Transport Processes and Associated Mathematical Modelling

5.1 Introduction

Sediments form a crucial link in estuarine processes. Suspended sediment concentrations are generally high, the particles are fine, cohesive, and prone to flocculate, and they are richly organic. The prediction of movements of very fine sediments is of major importance in many coastal areas or estuaries, as these sediments cause nuisances which require costly solutions: siltation in navigational channels and harbour docks, degradation of water quality. Cohesive sediments are also a vehicle for pollution, due to their faculty for absorption of heavy metals, pesticides and radionuclides (Teisson, 1991).
Large efforts have been undertaken to analyse extremely complex mechanism of transport, deposition and erosion of cohesive sediment as reported by Mehta (1986). So sediment transport in coastal and estuarine regions is an important topic which is receiving increasing attention by environmental researchers. The description of cohesive sediment transport process is given in section 5.1, and the review of some numerical modelling of cohesive sediment transport is given in section 5.2.

5.2 Cohesive Sediment Transport Process

Due to the interparticle forces, cohesive sediments and especially the very fine clay particles ($d < 2mm.$) form loose aggregates or flocs when the concentration of sediment in suspension increases, particularly in a saline environment. As flocs are formed, the settling velocity increases by several orders of magnitude compared to the individual clay particles. However, above a certain concentration, the particle aggregates start to hinder each other and the settling velocity decreases rapidly. When the flocs finally reach the bed, they form a dense fluid mud layer with strongly non-Newtonian rheological properties. Due to continuing deposition, the mud layer which initially was a loose fragile structure, gradually collapses under its increasing weight. The interstitial pore water is expelled; the weight of mud layer is progressively supported by the interparticle reaction forces. This process is called self-weight consolidation. The mud layer can be eroded due to fluid shear caused by currents or waves and induced turbulence.

In practice the prediction of cohesive sediment transport is intimately tied to the knowledge of these physical processes (Teisson, *et al.* 1993), which has encouraged extensive studies on these processes.

But cohesive sediments appear like a "water-sediment" complex, and the behaviour under flow action can be modified by the physico-chemical properties of the fluid (temperature, ionic composition,...) or by the sediment nature itself (mineralogic composition, organic content,...). Thus, for engineering applications, studies have usually been empirical and site specific, with the development of well known laws relating the rate of the sediment processes to "lumped" parameters (Mehta, 1989a; Delo 1988). These global parameters, such as bed density and mean flow velocity, do not account for the basic nature of particle-particle or flow-particle interaction (Mehta 1989a). This inhibited the intercomparison of data or even methodology, and the list of parameters to characterize mud. So there is a need to standardize the parameters, which aims at a better description of mud properties and understanding the basic cohesive sediment transport process.

Understanding the process of cohesive sediment, and quantifying them in terms of engineering parameters has, until recently, depended mainly on laboratory testing. Experiments focused on one isolated process (deposition, or erosion, or consolidation) and most usually under steady conditions. However, in marine environments, tides impose cyclic conditions and the processes of deposition at slack waters, partial consolidation, and resuspension at higher flow velocities are interrelated. Moreover, laboratory experiments cannot account for the complexity of the field. To enlarge the useful empirical

relationships derived from laboratory experiments for the processes in steady conditions (Mehta, 1988), there is a need to investigate some of the processes in unsteady conditions in laboratory experiments (Van Leussen and Winterwerp, 1990) or directly in the field (Van Leussen and Cornelisses, 1992).
Classical relationships describing the processes, although very useful, lack general applicability because the physics of the salient processes is hidden behind the "lumped" parameters. This is the case for instance for flow sediment interaction: classical soils theories applied to consolidation or two-phase flow models applied to turbulence show that, behind the diversity of behaviour or approaches, there must exist some firm relationships which can help in defining and then estimating the governing parameters.

5.2.1 Deposition

The process of deposition of cohesive sediment depends on a combination of different factors, including the size, settling velocity and strength of the settling particles. These particles may be single or, more likely, aggregates or flocs which may be loosely or strongly bound together. The flocs have dimensions and settling velocities of the primary particles. A review of the aggregation processes is given by Van Leussen (1988). A number of mechanisms can be responsible for the aggregation of the sediment particles, including salt flocculation, organic aggregation, bioflocculation, pelletization. Large fragile flocs may result, which are easily destroyed into smaller units, for example by shear force. It may be hypothesised that the deposition of the flocs is controlled by the stochastic turbulent processes in a zone near the bed. Only flocs that are strong enough to resist the bed shear stresses will settle to the bed. Flocs of which the strength is too low will be broken up into two or more smaller units and re-entrained into the suspension by the hydrodynamic lift forces.
Several researchers have formulated models for deposition, either for a uniform sediment or for a distributed sediment with different sized particles having a range of strengths and settling velocities (Mehta and Partheniades, 1975; Mehta and Lott, 1987; Verbeek *et al.* 1993; Krishnappan, 1991). Most of this work has been developed from laboratory tests where sediment is contained in a closed system. For application to field situations the models should be based on input also determined from the field. The field data recorded in the Mersey Estuary was used to test and compare a simple deposition model for uniform sediment and a second model for distributed sediment. The models were first test on laboratory data, and the sensitivity of the models to the input parameters was

also investigated (Ockenden, 1993).

For the simple uniform sediment model, deposition rate on to the bed, d, was modelled according to Krone (1962) by the equation:

$$d = (1 - \tau_b/\tau_d)w_s c, \qquad \tau_b \leq \tau_d. \tag{5.1}$$

Time histories of bed shear stress, τ_b, critical shear stress for deposition τ_d , settling velocity w_s and suspended sediment concentration, c, were prescribed during a tide. The model was then modified to allow for a distributed sediment according to Mehta and Lott (1987). This model assumes that the sediment is divided into N classes, each having a unique settling velocity, w_{si}, concentration, c_i and critical shear stress for deposition, τ_{di}. The total sediment deposited rate on the bed D, is given by the sum of the individual amounts deposited from each class:

$$D = \Sigma w_{si} f c_i c(1 - \tau_b/\tau_{di}) \tag{5.2}$$

where fc_i is the proportion of the total concentration in sediment class i, and sediment class i only deposits if $\tau_b \leq \tau_{di}$.

Both models gave a reasonable fit to the measured data for each of the monitored tides (Teisson, 1992). One noticeable feature (which occurs in each of the tides) is that the predicted deposition (from both models) occurs earlier than the measured deposition, showing a delay between the low shear stresses and the actual deposition on the bed. This delay is consistent with a flow-sediment hysteresis effect, as described by Costa and Mehta (1990), where for the same mean horizontal velocity, values of the Reynolds stress and turbulent variances were higher during decelerating flow than accelerating flow.

The distributed sediment model includes an improved representation of the differential settling of aggregates. However, for the application of field data, the uniform sediment model gives as good a fit to the data as the distributed sediment model.

The predicted deposition (and corresponding drop in concentration in suspension) over high water corresponds well with the measured drop in concentration recorded over this period (Ockenden, 1993), indicating that it is only a local effect. However, at other periods of the tide, it can be shown in the field data that sediment has been advected in from somewhere outside the region. By comparing the predicted and measured concentrations it should be possible to separate the local effects from the advected effects (Dyer, 1988).

5.2.2 Erosion

The mechanics of mud erosion has been reviewed by Mehta *et al.* (1989). Three modes of erosion have been identified according to the magnitude of the bed shear stress and the nature of the deposit. Visualisation of these processes with detailed picture of eroded beds can be found in (Migniot, 1968) and (Perigaud, 1983). This includes:

- Re-entrainment of stationary suspensions, where an undulation in the interface appears which is gradually accentuated, increasingly deforming the layer of sediment. In the end the sediment is carried away in the form of mud streaks and diluted in the current water.
- Surface erosion of consolidating beds where the eroded surface creases and the surface film is torn. The mud is eroded in the form of flakes which are more or less diluted. In this case, Migniot defined the mud as plastic.
- Mass erosion of fully consolidated deposits, which requires very large bottom velocities (1 m/s or more) to be eroded. Mud pebbles are formed and not easily diluted.

By analysing bed samples with a Brookfield viscometer, (Migniot, 1989) fixed the limit between fluid and plastic mud at a yield value of $3 N/m^2$ and that between plastic and fully consolidated muds at 75 N/m^2.
From the hydrodynamic point of view, bursts of turbulence near bed appear to play a dominant role in the erosion process (Winterwerp, 1989). Hydrodynamic forces are balanced by cohesion forces which are influenced by a large number of chemical and biological factors. There is no established theory for calculating the rate of erosion of mud deposits, instead it is necessary to rely upon site specific studies including laboratory flume tests or field measurements.
Many empirical laws for the erosion rate, E, have been proposed in the literature from deposited bed flume experiments. The simplest is expressed as (Ockenden *et al.* 1989):

$$E = M_e(\tau - \tau_c) \tag{5.3}$$

with M_e = a constant, τ = bed shear stress estimated in clear water, and τ_c = shear strength of the deposit.

M_e and τ_c are functions of the mud deposit and the depth. The determination of the critical shear stress τ_c is discussed in Berlamon *et al.* (1993). Based on the concept of floc erosion rate, Parchure and Mehta (1985) have proposed a more complex expression where two additional parameters are used including the floc erosion rate E_f:

$$E = E_f exp(a(\tau - \tau_c)^b) \qquad (5.4)$$

(Kujiper *et al.* 1989) have performed experiments in a recirculating straight flume and an annular flume with kaolinite and have analysed the results following Parchure's methodology. Two important conclusions have been drawn up from this study: (i) the larger number of parameters to be set facilitates the tuning of the equation to provide reasonable results. This tuning would be a cumbersome (or even unreliable task) for practical cases. (ii) the proposed formula is very sensitive to the bed shear stress which is very difficult to estimate in practical cases.

For practical applications, erosion formulae derived from steady state laboratory experiments are often used in tidal conditions. However, it is commonly observed in the field that during decreasing currents concentrations are higher than during increasing currents; likewise, sediment concentrations often lag hydrodynamic force (Costa and Mehta, 1990). This feature is known as flow-sediment hysteresis (Dyer, 1988).

Van Leussen and Winterwerp, (1990) conducted annular flume experiments under tidal conditions and showed the dominant role of the top layer of the bed, which develops under the successive phases of deposition and erosion. Costa and Mehta, (1990) also illustrated the difficulty of predicting suspended sediment concentration during a tidal cycle with a simple description of erosion/deposition processes that arose from laboratory experiments.

Long term simulation of suspended sediment transport in the Loire estuary performed by (Fritsch *et al.* 1989) has shown very little dependence on the erosion law, but the correct estimations of the critical shear stress was very important in this case.

5.2.3 Consolidation

During a deposition phase, flocs or individual particles settle on the bottom and form new sediments that can concentrate and consolidate. The stiffness of this fresh mud increases progressively and consequently its erodibility is reduced. Thus cohesive sediment transport modelling requires knowledge of the surficial sediment shear strength which has to be deduced from a consolidation model or measured in laboratory or field tests. A

parameter which quantifies the erodibility of a sediment is its erosion threshold. i.e. the bottom shear stress which is needed to resuspended it.
In the present work, investigations about self-weight consolidation have been conducted in four directions:

- synthesis of previous fundamental research, especially that drawn from soil mechanics;
- compilation of existing laboratory settling experiments and identification of required new tests;
- development of mathematical models for consolidation, and coupling with sediment transport models;
- validation with field measurements, both in terms of density evolution and shear strength variation.

As indicated in a review by Alexis *et al.* (1992), two families of models can be distinguished:

- The "sedimentation" models express mass conservation of the solid particles, with vertical exchanges represented by the settling fluxes. A common assumption of this approach is the unique dependence of the settling velocity on the local suspension density (Kynch, 1952), due to hindered settling processes;
- The so called "consolidation" models account for mass conservation of pore water and relate its expulsion between particles to the pressure vertical gradient by means of permeability, assuming Darcy's law. From the dynamics point of view, the stresses within the soil can be split into effective stresses on the grains and pore pressure on the fluid: only the latter forces the water movement. This concept is the basis of geotechnicians' approach. When combined with water mass conservation, it leads to Gibson's equation (Gibson *et al.* 1967), assuming constitutive relationships for permeability and effective stress as functions of the void ratio.

The former models seem to be more appropriate for suspensions and the latter convenient for dense mixtures; both are consistent as they consider the relative movement between the solid and fluid phases. Their analogy when effective stress is negligible has been shown

by Been, 1980. On the other hand, the possible reduction of settling velocity by effective stress can be introduced in a sedimentation model (Toorman and Berlamont, 1993). In fact, a real two phase flow approach should be more suitable when dealing with a large range of concentrations, enabling total continuity between deposition and consolidation processes.

Despite its theoretical basis, Gibson's equation does not represent a general formulation for consolidation, as it relies on restrictive assumptions. By not assuming Darcy's law and accounting for the fluid compressibility (partially filled voids for unsaturated soils. Alexis *et al.* (1992) proposed an extended formulation of the consolidation equation which is very similar to the settling formulation by Toorman and Berlamont, (1993).

It should be noted that all these theories do not deal with bed shear strength and require a relationship between concentration and erosion threshold for parameterizing the sediment erodibility. Besides, they only consider vertical processes, although some horizontal movement can interfere, especially when mud is fluidized by wave motion.

5.2.4 Turbulence

Turbulence is the factor which maintains sediment in suspension by opposing the settling flux. The vertical distribution of sediment results from these antagonistic actions and has been thoroughly described under varying hydrodynamic forcing (Mehta, 1989b). However, our knowledge is still hampered by the complex interactions between the sediment and the flow field, especially at large concentrations or near the bottom.

Some attempts have already been made to account for the influence of sediment in these two areas:

(i) The vertical structure of the concentration profile in highly concentrated areas and the presence of so-called lutoclines — zones of high gradient of concentration and minimum of mixing — has been described and mathematically modelled by (Wolanski *et al.* 1988; Ross and Mehta, 1989; and Smith and Kirby, 1989). In these models, sediment is mixed vertically by an eddy diffusivity approach. However, through a Richardson number dependency, sediment-induced density effects inhibit vertical mixing. These models enable satisfactory reproduction of the generation and evolution of lutoclines.

(ii) The other domain of interest is how the presence of sediment can affect the bottom shear stress and the bottom processes. (Sheng and Villaret, 1989) performed numerical test with a turbulent 1DV model including investigation of the influence of the suspension on the turbulent flow and bed shear stresses. This model has been applied further

by (Huynh-Thanh *et al.* 1991). Gust, (1976) verified the turbulent draft reduction in a clay suspension in a small flume and found in the field that the shear velocity u^* could be reduced by as much as 40%.

Classically, the variation of sediment concentration c along the vertical axis is written as:

$$\frac{\partial \bar{c}}{\partial t} + \frac{\partial w_s \bar{c}}{\partial z} = -\frac{\partial \overline{w'c'}}{\partial z} \tag{5.5}$$

where w is the vertical velocity of the flow, the overbar sign is the time-average value, w_s is the settling velocity of particles, possibly a function of concentration and c', w' are the fluctuation of $\bar{c}$ and $\bar{w}$ respectively.

The turbulent flux is usually modelled by a first gradient approximation:

$$\overline{w'c'} = -K \partial \bar{c} / \partial z \tag{5.6}$$

and the effect of concentration (density) on turbulence is generally not represented, leading in steady condition to the well known Rouse equilibrium profiles.

For the boundary condition at the bottom, exchange with the bed (erosion and sedimentation) is empirically related to a mean flow velocity.

In low turbidity environments, these models are acceptable and have been widely applied, but they present two inadequacies for a more fundamental approach: to rely on a crude approximation of turbulence, and to relate processes occurring near the bed to lumped parameters. In essence, they cannot represent accurately the processes very close to the bottom when large concentrations are experienced, and flow sediment interaction becomes important. Neither can they describe thoroughly the balance or imbalance between turbulent and settling fluxes. Consequently, laboratory studies have been concerned with providing deposition and erosion rate expression in relation with the level of mathematical modelling, with less than adequate emphasis placed on the evolution of the vertical structure (Mehta, 1988). Application of classical erosion/deposition laws in high concentration environments has revealed itself of limited utility due to complex near bed interactions and has suggested the need for an improvement in turbulence modelling (Costa and Mehta, 1990).

Today, a new generation of models , such as Reynolds stress models and two-phase flow (sediment and fluid) models has become available in the industrial domain experts have tried to apply and the most recent ideas and concepts to cohesive sediment laden flow. The philosophy is to represent the hydrodynamics, turbulence and flow sediment interaction in the most accurate way, and, if possible, to get a new understanding of the physical processes in return.

5.3 Numerical Modelling

The implementation of these processes into numerical models has started in the 1970's with Odd and Owen, 1972 and Ariathurai, 1974. At that time, use of three-dimensional (3D) models was impossible due to computing costs and capacity: morphological evolutions are very slow in the field and require time consuming long term simulation.
Several numerical models have been developed in recent years. These models range from one-dimensional vertical profile formulations (e.g. De Vantier and Narayanaswamy, 1989; Hagatun and Eidsvik, 1986) via two-dimensional vertical or depth integrated models (e.g. Van Rijn, 1987 and references therein; Teisson and Frisch 1988; Veeramachaneni and Hayter, 1988) to quasi- or full three-dimensional formulations (Sheng and Bulter, 1982; Van Rijn and Meijer, 1988; O'Connor and Nicholson, 1988; Eidsvik and Utnes 1991; Utnes 1993;Utnes and Ren, 1995) and others.
For engineering applications, the sediment transport model is generally coupled with a hydrodynamic model, to get information on flow velocities. All above mentioned models are based on the conventional Reynolds averaging of the incompressible Navier-Stokes equation including the continuity equation, turbulence model and advection-diffusion concentration equation. The model equations are obtained by a conventional Reynolds averaging of the incompressible Navier-Stokes equations including the continuity equation. The governing equations of some models will be given later.
The equations which govern the distribution of the mean flow quantities are dealt with in this section and therefore form the basis of the so-called field methods. The origin of these equations are the conservation laws for mass, momentum and suspended sediment concentration. For incompressible flows, these laws can be expressed in tensor notation as follows. A short introduction to tensor notation is given as Appendix 5.A.

Mass conservation: continuity equation:

$$\frac{\partial U_i}{\partial x_i} = 0, \qquad i = 1,2,3 \tag{5.7}$$

Momentum conservation: Navier-Stokes equations:

$$\frac{\partial U_i}{\partial t} + U_j \frac{\partial}{\partial x_j}(U_i) = -\frac{1}{\rho}\frac{\partial P}{\partial x_i} + \nu \frac{\partial^2 U_i}{\partial x_i \partial x_j} + g_i, \qquad i,j = 1,2,3 \tag{5.8}$$

Sediment concentration conservation:

$$\frac{\partial C}{\partial t} + U_i \frac{\partial C}{\partial x_i} = \lambda \frac{\partial^2 C}{\partial x_i \partial x_j} + S_c \tag{5.9}$$

Where U_i are the instantaneous velocity component in the direction x_i, P is the instantaneous static pressure, C is the instantaneous suspended sediment concentration, S_c is a source term expressing the mass generation due to chemical or biological reactions, ν, λ are the kinematic viscosity and diffusivity of C respectively, ρ is the flow density that is assumed as constant and g_i is the gravitational acceleration in direction x_i.

Reynolds' time-averaging procedure

Reynolds's applied the Navier-Stokes equations to turbulent flow by introducing a time-averaging procedure. Each instantaneous variable is represented as a time-mean value plus a fluctuating value. Thus:

$$U_i = u_i + u_i', \quad i = 1,2,3 \quad C = c + c', \qquad P = p + p' \tag{5.10}$$

u_i are the time-mean velocity components in the direction x_i, and u_i' are the fluctuation velocity component in the direction of x_i. p, p' are the time-mean and fluctuation static pressure respectively and c, c' are the time-mean and fluctuation suspended concentration respectively.
The mean values are defined by:

$$f = \frac{1}{T} \int_0^T F(t) dt \tag{5.11}$$

in which T is time-averaging period. This period T should be larger than the dominant turbulence scale, but small than the long periodic effects such as the tidal scale (T between 2 to 5 minutes is a good choice).

5.3.1 Three Dimensional Model

The governing 3D model equations are usually obtained by a conventional Reynolds averaging of the incompressible Navier-Stokes equations including the continuity equation as follows:

Momentum Equations:

$$\frac{\partial u_i}{\partial t} + \frac{\partial}{\partial x_j}(u_i u_j) = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}(\nu \frac{\partial u_i}{\partial x_j} - \overline{u_i' u_j'}) + g_i, \qquad i,j = 1,2,3. \tag{5.12}$$

Continuity equation:

$$\frac{\partial u_i}{\partial x_i} = 0, \qquad i = 1,2,3 \tag{5.13}$$

The suspended sediment concentration is computed from the transport equation as follows:

$$\frac{\partial c}{\partial t} + \frac{\partial}{\partial x_i}(c u_i) = \frac{\partial}{\partial x_i}(\frac{\nu_t}{\sigma_c}\frac{\partial c}{\partial x_i} - \overline{u_i' c}) + \frac{\partial c w_s}{\partial x_i} + S_c \tag{5.14}$$

Here the overbar sign means the time average value and $\overline{u_i' u_j'}$ is the turbulent Reynolds stress tensor.

The Reynolds stress is modelled by use of the Boussinesq assumption

$$\overline{u_i' u_j'} = \nu_t(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}) - \frac{2}{3}k\delta_{ij} \tag{5.15}$$

where

$$\nu_t = C_\mu \frac{k^2}{\epsilon} \tag{5.16}$$

is the eddy viscosity,

$$k = \frac{1}{2}(\overline{u_1'^2} + \overline{u_2'^2} + \overline{u_3'^2}) \tag{5.17}$$

is the mean turbulent kinetic energy, ϵ is the turbulent energy dissipation rate and δ_{ij} is the Kronecker delta.

These are the equations governing the mean-flow quantities u_i, p and c. The equations are also exact since no assumptions have been introduced in deriving them; but they no longer form a closed set due to the nonlinearity of (5.8) and (5.9). The averaging process has introduced unknown correlations between fluctuating velocities, $\overline{u_i' u_j'}$, and scalar fluctuations $\overline{u_i' c}$. Physically, these correlations, multiplied by the density ρ, represent the transport of momentum and mass due to fluctuating (i.e. turbulent) motion. $-\rho\overline{u_i' u_j'}$ is the transport of x_i- momentum in the direction x_j (or vice versa); it acts as a stress on the fluid and is therefore called turbulent or Reynolds stress. $-\rho\overline{u_i' c}$ is the transport of the

scalar quantity c in the direction x_i and is therefore a turbulent mass flux. In most flow regions, the turbulent stresses and fluxes are much larger than their laminar counterparts $\nu\frac{\partial u_i}{\partial x_j}$ and $\lambda\frac{\partial c}{\partial x_j}$ which are often negligible (Rodi, 1979).
Equations (5.12)-(5.14) can be solved for the mean values of velocity, pressure and sediment concentration only when the turbulence correlations $\overline{u_i'u_j'}$ and $\overline{u_i'c}$ can be determined in some way. In fact, the determination of these correlations is the main problem in calculating turbulence flows. Exact transport equations can be derived for $\overline{u_i'u_j'}$ and $\overline{u_i'c}$ that will be given later, but these equations contain turbulence correlations of the next higher order. Therefore, closure of the equations cannot be obtained by resorting to equations for correlations of higher and higher order; instead, a turbulence model must be introduced which approximates the correlations of a certain order in terms of lower order correlations and mean-flow quantities. The laws described by a turbulence model simulate the averaged character of real turbulence; these laws are expressed in differential and algebraic equations which together with the mean flow equation (5.12) to (5.14) form a closed set.

5.3.2 The Standard High Reynolds Number $(k-\epsilon)$ Model

$$\underbrace{\frac{\partial k}{\partial t}}_{\text{rate of change}} + \underbrace{u_i\frac{\partial k}{\partial x_i}}_{\text{convective transport}} = \underbrace{\frac{\partial}{\partial x_i}(\frac{\nu_t}{\sigma_k}\frac{\partial k}{\partial x_i})}_{\text{diffusive transport}} + \underbrace{\nu_t(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i})\frac{\partial u_i}{\partial x_j}}_{P=\text{production by shear / destruction}} + \underbrace{\beta g_i\frac{\nu_t}{\sigma_t}\frac{\partial c}{\partial x_i}}_{G=\text{buoyant production}} - \underbrace{\nu\overline{\frac{\partial u_i}{\partial x_j}}\overline{\frac{\partial u_i}{\partial x_j}}}_{\epsilon=\text{viscous dissipation}} \tag{5.18}$$

$$\underbrace{\frac{\partial \epsilon}{\partial t}}_{\text{rate of change}} + \underbrace{U_i\frac{\partial \epsilon}{\partial x_i}}_{\text{convection}} = \underbrace{\frac{\partial}{\partial x_i}(\frac{\nu_t}{\sigma_\epsilon}\frac{\partial \epsilon}{\partial x_i})}_{\text{diffusion}} + \underbrace{c_{1\epsilon}\frac{\epsilon}{k}(P+G)(1+c_{3\epsilon}R_f) - c_{2\epsilon}\frac{\epsilon^2}{k}}_{\text{generation-destruction}} \tag{5.19}$$

and the conventional model constants are (Rodi, 1979):

$$c_{1\epsilon}=1.44, \quad c_{2\epsilon}=1.92, \quad c_{3\epsilon}=0.8 \quad \sigma_k=1.0, \quad \sigma_\epsilon=1.3$$
$$\sigma_c=1.0 \quad \beta=10, \quad R_f=-G/(P+G),$$

The deposition, erosion and consolidation information are usually included in the boundary conditions.

5.3.3 Reynolds-Stress Equations

The first suggestion to determine $\overline{u_i'u_j'}$ from a transport equation was made already in (Keller and Friedmann, 1924). These authors showed how (under certain assumptions) equations for $\overline{u_i'u_j'}$ can be derived, but they did not give the equations explicitly. (Chou, 1945) was the first to derive and present the exact $\overline{u_i'u_j'}$ equation given below:

$$\underbrace{\frac{\partial \overline{u_i'u_j'}}{\partial t}}_{\text{rate of change}} + \underbrace{\overline{u_l'}\frac{\partial \overline{u_i'u_j'}}{\partial x_l}}_{\text{convective transport}} = \underbrace{-\frac{\partial}{\partial x_l}(\overline{u_l'u_i'u_j'}) - \frac{1}{\rho}(\frac{\partial \overline{u_j'p}}{\partial x_i} + \frac{\partial \overline{u_i'p}}{\partial x_j})}_{\text{diffusive transport}}$$
$$- \underbrace{\overline{u_i'u_l'}\frac{\partial u_j}{\partial x_l} - \overline{u_j'u_l'}\frac{\partial u_i}{\partial x_l}}_{P_{ij}=\text{stress production}} - \underbrace{\beta(g_i\overline{u_j'c} + g_j\overline{u_i'c})}_{G_{ij}=\text{buoyancy production}}$$
$$+ \underbrace{\overline{\frac{p}{\rho}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i})}}_{\pi_{ij}=\text{pressure strain}} \quad \underbrace{-2\nu\overline{\frac{\partial u_i}{\partial x_l}\frac{\partial u_j}{\partial x_l}}}_{\epsilon_{ij}=\text{viscous dissipation}} \tag{5.20}$$

The contraction of this equation, that is when the 3 equations for the 3 normal stresses ($i = j = 1, 2, 3$) are assumed up, yields the exact turbulent kinetic energy equation (5.18) presented already (note that $k = \frac{1}{2}\overline{u_i'^2u_i'^2}$). The physical meaning of the individual terms of the k-equation was described and equivalent terms appear in equation (5.20) which represent the rate of change, convective and diffusive transport, stress- and buoyancy-production, and viscous destruction of $\overline{u_i'u_j'}$. Equation (5.20) contains an additional term denoted "pressure-strain" term because it involves correlation between fluctuating pressure and strain rates. This term is absent in the k-equation (5.18) so that it contributes nothing to the total energy balance; it only acts to redistribute the energy among its components (when $i = j$) and to reduce shear stresses (when $i \neq j$). This term tends therefore to make the turbulence more isotropic.

5.3.4 Two Dimensional Depth-integrated Model

In case of flow in a large-scale area in which the water depth is small compared with the horizontal dimensions of the area concerned, it is allowed to neglect the accelerations in the vertical plane. This leads to $\partial p/\partial z = -\rho g$, giving a hydrostatic pressure distribution ($p = -\rho g z$). A definition sketch is given in Fig 5.1.

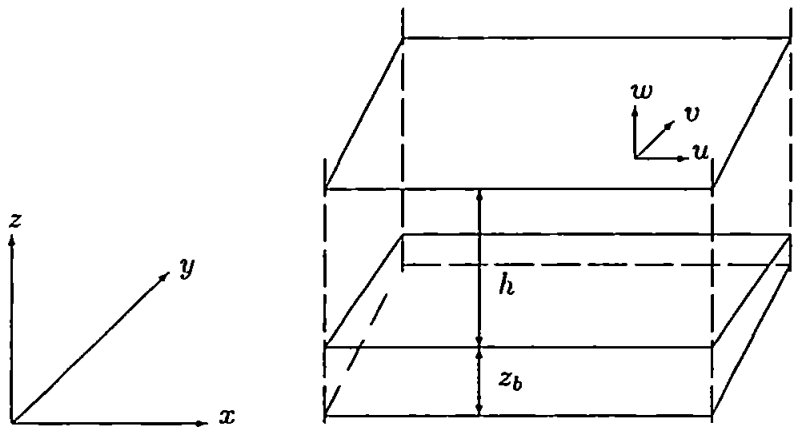


Fig 5.1 Symbol definitions

Depth-integrated models therefore appeared very attractive in well mixed situations, and numerous two-dimensional horizontal (2DH) models were developed, for example are those of (Hayter and Mehta 1982, Cole and Miles 1983, Thomas and McAnally 1985, and Teisson and Latteux 1986). All these models are based on a time-average continuity equation, Reynolds stresses equation, $k-\epsilon$ turbulence model and the well known advection-diffusion equation. The equations of motion in x and y direction can be integrated over the water depth, yielding (ρ = constant):

Continuity equation:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(h\overline{u}) + \frac{\partial}{\partial y}(h\overline{v}) = 0 \tag{5.21}$$

Momentum equations:

$$\begin{aligned}\frac{\partial h\overline{u}}{\partial t} &+ \frac{\partial h\overline{u}^2}{\partial x} + \frac{\partial h\overline{u}\,\overline{v}}{\partial y} + gh\frac{\partial(h+z_b)}{\partial x} + \frac{1}{\rho}\tau_{bx} \\ &- \frac{1}{\rho}\frac{\partial}{\partial x}(h\overline{\sigma}_{xx} + hK_{xx}) - \frac{1}{\rho}\frac{\partial}{\partial y}(h\overline{\tau}_{xy} + hK_{xy}) - \frac{1}{\rho}F_{c,x} - \frac{1}{\rho}\Sigma F_x = 0\end{aligned} \tag{5.22}$$

$$\frac{\partial h\overline{v}}{\partial t} + \frac{\partial h\overline{v}^2}{\partial y} + \frac{\partial h\overline{u}\,\overline{v}}{\partial x} + gh\frac{\partial(h+z_b)}{\partial y} + \frac{1}{\rho}\tau_{by} - \frac{1}{\rho}\frac{\partial}{\partial y}(h\overline{\sigma}_{yy} + hK_{yy}) - \frac{1}{\rho}\frac{\partial}{\partial x}(h\overline{\tau}_{xy} + hK_{xy}) - \frac{1}{\rho}F_{c,y} - \frac{1}{\rho}\Sigma F_y = 0 \qquad (5.23)$$

Advection-diffusion equation:

$$\frac{\partial\overline{c}}{\partial t} + u\frac{\partial\overline{c}}{\partial x} + v\frac{\partial\overline{c}}{\partial y} = \frac{1}{h}\frac{\partial}{\partial x}[hK_x\frac{\partial\overline{c}}{\partial x}] + \frac{1}{h}\frac{\partial}{\partial y}[hK_y\frac{\partial\overline{c}}{\partial y}] + \frac{S}{h} \qquad (5.24)$$

in which

h = water depth

S = source-sink term

$\overline{c}$ $= \frac{1}{h}\int_{z_b}^{z_s} c\,dz$ = depth-integrated concentration of suspended load

$\overline{u}$ $= \frac{1}{h}\int_{z_b}^{z_s} u\,dz$ = depth-averaged velocity in x-direction.

$\overline{v}$ $= \frac{1}{h}\int_{z_b}^{z_s} v\,dz$ = depth-averaged velocity in y-direction.

z_s = vertical coordinate of water surface above a horizontal plane

z_b = vertical co-ordinate of bottom above a horizontal plane

F_{cx}, F_{cy} = body force per unit area due the earth rotation (coriolis force) in x and y direction

F_x, F_y = external driving forces (wave-induced, wind induced) per unit area

τ_{bx}, τ_{by} = bottom shear stress in x-direction and y- direction

$\overline{\tau}_{xy}$ $= \frac{1}{h}\int_{z_b}^{z_s}\{2\rho\nu(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}) - \rho\overline{u'v'}\}dz$ = depth-average shear stress

$\overline{\sigma}_{xx}$ $= \frac{1}{h}\int_{z_b}^{z_s}\{2\rho\nu\frac{\partial u}{\partial x} - \rho\overline{u'u'}\}dz$ = depth-average normal stress

$\overline{\sigma}_{yy}$ $= \frac{1}{h}\int_{z_b}^{z_s}\{2\rho\nu\frac{\partial v}{\partial y} - \rho\overline{v'v'}\}dz$ = depth-average normal stress

K_{xy} $= -\frac{1}{h}\int_{z_b}^{z_s}\rho(\overline{u}-u)(\overline{v}-v)dz$ = dispersion coefficient

K_{xx} $= -\frac{1}{h}\int_{z_b}^{z_s}\rho(\overline{u}-u)(\overline{u}-u)dz$ = dispersion coefficient

K_{yy} $= -\frac{1}{h}\int_{z_b}^{z_s}\rho(\overline{v}-v)(\overline{v}-v)dz$ = dispersion coefficient

K_x, K_y = effective dispersion coefficient

Usually, the depth-averaged stresses and the dispersion components (introduced by the depth-averaging procedure) are related to local velocity gradients as following:

$$\frac{\partial}{\partial x}(h\overline{\sigma}_{xx} + hK_{xx}) + \frac{\partial}{\partial y}(h\overline{\tau}_{xy} + hK_{xy}) = K_x(\frac{\partial^2\overline{u}}{\partial x^2} + \frac{\partial^2\overline{v}}{\partial y^2}) \qquad (5.25)$$

$$\frac{\partial}{\partial y}(h\overline{\sigma}_{yy} + hK_{yy}) + \frac{\partial}{\partial x}(h\overline{\tau}_{xy} + hK_{xy}) = K_y(\frac{\partial^2\overline{v}}{\partial x^2} + \frac{\partial^2\overline{u}}{\partial y^2}) \qquad (5.26)$$

$k-\epsilon$ model for depth-average calculations

It is assumed that the local depth-average state of turbulence can be characterised by the energy and dissipation $\tilde{k}$ and $\tilde{\epsilon}$. The variation of $\tilde{k}$ and $\tilde{\epsilon}$ is determined from the following transport equations:

$$\overline{u}\frac{\partial \tilde{k}}{\partial x} + \overline{v}\frac{\partial \tilde{k}}{\partial y} = \frac{\partial}{\partial x}(\frac{\tilde{\nu}_t}{\sigma_k}\frac{\partial \tilde{k}}{\partial x}) + \frac{\partial}{\partial y}(\frac{\tilde{\nu}_t}{\sigma_k}\frac{\partial \tilde{k}}{\partial y}) + P_h + P_{kv} - \tilde{\epsilon} \tag{5.27}$$

$$\overline{u}\frac{\partial \tilde{\epsilon}}{\partial x} + \overline{v}\frac{\partial \tilde{\epsilon}}{\partial y} = \frac{\partial}{\partial x}(\frac{\tilde{\nu}_t}{\sigma_\epsilon}\frac{\partial \tilde{\epsilon}}{\partial x}) + \frac{\partial}{\partial y}(\frac{\tilde{\nu}_t}{\sigma_\epsilon}\frac{\partial \tilde{\epsilon}}{\partial y}) + c_{1\epsilon}\frac{\tilde{\epsilon}}{\tilde{k}}P_h + P_{\epsilon v} - c_{2\epsilon}\frac{\tilde{\epsilon}^2}{\tilde{k}} \tag{5.28}$$

where

$$P_h = \tilde{\nu}_t[2(\frac{\partial \overline{u}}{\partial x})^2 + 2(\frac{\partial \overline{v}}{\partial y})^2 + (\frac{\partial \overline{u}}{\partial y} + \frac{\partial \overline{v}}{\partial x})^2] \tag{5.29}$$

is the production of $\tilde{k}$ due to interactions of turbulent stresses with horizontal mean velocity gradients. $\tilde{k}$, $\tilde{\epsilon}$ and $\tilde{\nu}_t$ are not strictly depth-average values (they are defined by the above equations), but the $\tilde{k}$- and $\tilde{\epsilon}$- equation (5.27) and (5.28) can still be considered as depth-averaged forms of three-dimensional equations (5.18) and (5.19) as can all terms originating from non-uniformity of vertical profiles which are assumed to be absorbed in the source term P_{kv} and $P_{\epsilon v}$. The main contribution to these terms stems from significant vertical velocity gradients near the bottom of the water body. By interaction with the relatively large turbulent shear stresses in this region these gradient produce turbulence energy, which is in addition to the production P_h due to horizontal velocity gradients and which depends strongly on the bottom roughness. Rastogi and Rodi, (1978) related the additional source term P_{kv} and $P_{\epsilon v}$ to the friction velocity u^* by writing

$$P_{kv} = c_k\frac{u^{*3}}{h}, \qquad P_{\epsilon v} = c_\epsilon\frac{u^{*4}}{h^2} \tag{5.30}$$

and determined u^* from the usual quadratic friction law

$$u^* = \sqrt{c_f(\overline{u}^2 + \overline{v}^2)} \tag{5.31}$$

where c_f is a friction coefficient. The empirical constants c_k and c_ϵ were determined from undisturbed normal channel flow

$$c_k = \frac{1}{\sqrt{c_f}}, \qquad c_\epsilon = 3.6\frac{c_{2\epsilon}}{c_f^{3/4}}\sqrt{c_\mu} \tag{5.32}$$

The adaptation of the $k-\epsilon$ model for depth-averaged calculations is certainly of a rather empirical nature, but the calculations performed so far (Rastogi and Rodi, 1978; McGuirk and Rodi, 1978; and Hassain, 1980) show encouraging results; It should be emphasised that the model described here does not account for the dispersion terms appearing in equation (5.22) to (5.24).

5.3.5 One Dimensional Two-phase Flow Model

It is assumed that the sediment concentration c is small so that the interaction effects between particles are negligible. The response time of the particles is assumed to be small compared to the characteristic time to the mean flow so that, except for a systematic fall velocity, the particles follow the mean flow. The waves and currents are mainly oriented along one direction characterised by the co-ordinate x. With these simplifications, the conservation equations for mean momentum and suspended mass become:

$$\frac{\partial \overline{\rho}\,\overline{u}}{\partial t} = -\frac{\partial \overline{p}}{\partial x} + \frac{\partial}{\partial z}[\overline{\rho}(\nu_T + \nu_L)\frac{\partial \overline{u}}{\partial z}] \tag{5.33}$$

$$\frac{\partial k}{\partial t} = \frac{\partial}{\partial z}[(\nu_T/\sigma_k + \nu_L)\frac{\partial k}{\partial z}] + P + G - \epsilon \tag{5.34}$$

$$\frac{\partial \epsilon}{\partial t} = \frac{\partial}{\partial z}[(\nu_T/\sigma_\epsilon + \nu_L)\frac{\partial \epsilon}{\partial z}] + [c_1(P+G) - c_2\epsilon]\frac{\epsilon}{k} \tag{5.35}$$

$$\frac{\partial \overline{c}}{\partial t} = \frac{\partial}{\partial z}[\overline{c}w_s + (\frac{\nu_T}{\sigma_c} + \nu_L)\frac{\partial \overline{c}}{\partial z}] \tag{5.36}$$

Here the force function is the pressure gradient $\partial\overline{p}/\partial x$. The mass average density is $\overline{\rho} = (1-\overline{c})\rho_f + \overline{c}\rho_s$ and ρ_f, ρ_s are the density of sediment and flow respectively and $\Delta\rho = \rho_s - \rho_f$ is the density difference between the sediment and the flow. The mass average velocity $\overline{u}$ is given by $\overline{\rho u} = (1-\overline{c})\rho_f u_f + \overline{c}\rho_s u_s$. The settling velocity $w_s \propto (1-\overline{c})^n, n \approx 4$ (Sleath, 1984), ν_T, ν_L are the turbulent and molecular viscosity respectively, k is the turbulent kinetic energy, ϵ is the turbulent dissipation, $P = \nu_T(\partial\overline{u}/\partial z)^2$ is the mechanical energy production, $G = g(\Delta\rho/\rho)(\nu_T/\sigma_c)\partial\overline{c}/\partial z$ is the buoyant energy prodution, and $c_1, c_2, \sigma_c, \sigma_\epsilon, \sigma_k$ are parameters (Rodi 1979). At the free-surface, the net flux of suspended sediments are assumed to be zero. At the bottom, the net upward flux of suspended sediments is the difference of erosion rate E and the deposition rate D.

$$w_s c - (\frac{\nu_T}{\sigma_c} + \nu_L)\frac{\partial \overline{c}}{\partial z} = E - D \tag{5.37}$$

The particular forms of erosion and deposition rates are based on extensive literature survey and model testing with the field and laboratory data.

Various numerical models mentioned above, have been developed over the last decade to reproduce cohesive sediment transport. At the moment, these models supply interesting but only qualitative results; predictive results are far from being reached and the use

of models is often restricted merely to sensitivity analysis. This relative failure in gaining quantitative results is not due to the numerical techniques but rather to the incomplete knowledge of basic processes such as deposition, erosion and consolidation. Reliability in model predictions will increase with a better insight into these physical processes.
The unsatisfactory predictions of numerical models may also come from a possible discrepancy between specifications of physical laws, issued of laboratory experiments, and prototype behaviour. There is a trend to investigate in the field factors controlling physical processes, namely formation and size distribution of mud flocs (Van Leussen, 1988), deposition (Delo, 1988). Erosion, consolidation, wave effects are also studied in the field (Teisson, 1991).

In our view, the cohesive sediment dynamics is a very complicated system which depends on the site, season, weather condition and a lot of processes which are still not very clearly known and are rapidly changing. Also since a lot of factors that effect the cohesive sediment dynamics are very stochastic, we should pay more attention in the real field data to use the on-line modelling to trace the sediment dynamics and try to find some suitable models for it. This is why we use the time series model to try to find the answer in the cohesive sediment dynamics based on the real data collecting and give more accurate description and prediction of it. The background and application of time series models is given in earlier chapters and the description and prediction the mean flow quantities are given in chapters 6 and 7. It will be shown that this kind of model is good for the description, prediction and adaptation in sediment dynamics based on the *in situ* data collected.

Appendix 5.A. Introduction to Tensor Notation

Tensor notation is used in this chapter because it allows most equation to be written in a considerably more compact form than is possible with the conventional notation. A short introduction is provided as follows.

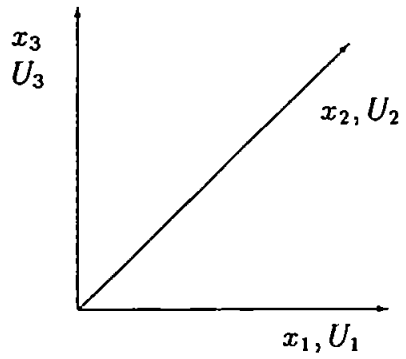


Fig 5.A.1. Cartesian coordinate system

In cartesian tensor notation, vector quantities are written by attaching an index to the symbol denoting the quantity, for example

space vector	$x_i \equiv \{x_1, x_2, x_3\}$
velocity vector	$U_i \equiv \{U_1, U_2, U_3\}$

The three components of the vector in a cartesian system (Figure 5.A.1.) are obtained by setting the index (here i) equal to 1,2, and 3, respectively.

A quantity with 2 indices (e.g. i and j) is called a tensor and has 9 components which can be obtained by pemutation of the 2 indices from 1 to 3:

$$a_{ij} \equiv \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

The stresses τ_{ij} appearing in the momentum equations (5.22) and (5.23) are an example of such a tensor; here the first index denotes the surface ($\perp$ to x_i) on which the stress acts and the second index the direction of the stress. The (diadic) product of two velocity vectors also yields a tensor:

$$U_iU_j \equiv \begin{pmatrix} U_1U_1 & U_1U_2 & U_1U_3 \\ U_2U_1 & U_2U_2 & U_2U_3 \\ U_3U_1 & U_3U_2 & U_3U_3 \end{pmatrix}$$

When the U's are the fluctuating velocities and the average of the products is taken (indicated by an overbar), this is the Reynolds stress $\overline{u_i'u_j'}$ tensor introduced in (5.15). The

tensors appearing in this chapter are all symmetric, that is $\tau_{ij} = \tau_{ji}$ and $\overline{u_i' u_j'} = \overline{u_j' u_i'}$ so that they have only 6 different components.

A particular tensor is the Kronecker delta δ_{ij} which has the components

$$\delta_{ij} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

so that $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$.

One aspect of tensor notation remains to be explained which is particularly effective in making equations more compact. This is the summation convention which implies that whenever the same index is repeated in a single expression, the sum over all 3 directions has to be taken, thus

$$U_i U_i = \sum_{i=1}^{3} U_i U_i = U_1 U_1 + U_2 U_2 + U_3 U_3$$

The continuity equation (5.7) may be cited here as further example:

$$\frac{\partial U_i}{\partial x_i} = \sum_{i=1}^{3} \frac{\partial U_i}{\partial x_i} = \frac{\partial U_1}{\partial x_1} + \frac{\partial U_2}{\partial x_2} + \frac{\partial U_3}{\partial x_3} = 0$$

and the second term on the left hand side of the momentum equation (5.8) reads in full:

$$U_j \frac{\partial U_i}{\partial x_j} = \sum_{j=1}^{3} U_j \frac{\partial U_i}{\partial x_j} = U_1 \frac{\partial U_i}{\partial x_1} + U_2 \frac{\partial U_i}{\partial x_2} + U_3 \frac{\partial U_i}{\partial x_3} \tag{5.38}$$

Chapter 6

Time Series Modelling for Sediment Dynamics On the Holderness Coast

6.1 Introduction

Sediment transport in the coastal region is a topic which is receiving increased attention due to its importance to flow environment. Among the main problems is the transport of pollutants which adhere to suspended sediments. In order to gain more knowledge of such process, several numerical models have been developed in recent years that try to predict the transport of sediment. These models range from one dimensional vertical profile formulations (e.g. De Vantier and Narayanaswamy, 1989; Hagagun and Eidsvik, 1986) via two dimensional vertical or depth integrated models (e.g. Van Rijn, 1987 and refs therein; Teisson and Frisch, 1988) to quasi or fully three dimensional formulations (e.g. Sheng and Bulter, 1982; Van Rijn and Meijer, 1988; O'Connor and Nicholson, 1988; Eidsvik and Utnes, 1991; Utnes, 1993 and others).

In the above mentioned papers, the Navier-Stokes Equation and its average form, turbulence closure models combined with some physical, chemical and rheological properties are used. But as is well known, sediment transport is a very complicated process. Within the cohesive sediment transport process, there are many physical, geographical, chemical and microbiological processes that are still not very well understood (Dyer 1989).

Recently some papers have been published that study ocean models and water movement in unsaturated soils used the data to estimate some parameters of the model based on the concept of inverse modelling. Copeland *et al.* (1991) presented a mathematical box model of a theoretical 'ocean' in which the field of water velocity, turbulence and tracer concentrations are all known and together 'perfectly' satisfy a steady-state advective-diffusive equation. Mous (1993) used nonlinear regression techniques to estimate the unknown parameter in his model but it appears to be non-identifiable, which results in non-unique solutions. Bagchi and ten Brummelhuis (1996) applied parameter identification technique to tidal models with uncertain boundaries and they considered a simultaneous state and parameter estimation procedure for tidal models with random inputs, which is formulated as a minimization problem. There is, however, no tradition for our type of time series modelling in sediment dynamics.

In this chapter, stochastic time series models are set up to describe the concentration of SPM and current velocity respectively. The recursive least squares identification algorithm is used to identify the unknown parameters of the model. The simulations are given to show the good approximation of the *in situ* data collected from the Holderness Coast by Joanna Blewett, Institute of Marine Science, University of Plymouth.

The aims of the chapter are:
(i) To describe and prove the accuracy of the time series model for current velocity and suspended sediment concentration dynamical system based on the data from the field.
(ii) To use the model to predict the sediment transport dymamics

6.2 Site Description and Data Collecting

The coastline of Holderness Cliffs which extend 61.5Km from Flamborough Head in the North to the sand and shingle spit of Spurn Head in the south, takes the plan shape of a zeta curve (Pringle, 1985). The cliffs are largely composed of Pleistocene glacial till and are renowned for their rapid rate of erosion, calculated by Pringle (1985) and Hardisty (1986) at 1.7 and 1.75myr-1 respectively. The result is that one million m^2 of high quality agricultural land is lost to the North Sea each decade. The cliff material provides sand and coarser sediment for beach replenishment whilst the silt and clay is carried away in suspension by the waves and currents. Maximum cliff and beach erosion is thought to

occur during storm events, but the magnitude of the storm largely controls the amount of material being washed offshore.

In order to test the time series models, point measured observations of the appropriate parameters, i.e. current velocity, suspended sediment concentration and pressure variation are required over a time scale of weeks. The data set chosen to test the models is taken from 29th June to 15 July, 1995, at the inner site (N1, Figure 6.2.1). Figure 6.2.2 shows the resultant time series of SPM concentrations taken at 0.4m above the seabed by the lower OBS sensor, longshore current velocity component, again at 0.4m above the seabed, and pressure variance which indicates the presence of waves and storms, all the data were used as input parameters to test the models. The average water depth at the inner station is about 10 meters. A grab sample taken just before the deployment, revealed that the surface sediment comprised of fine silt clay material overlaying more coarse sand.

During the time series (Figure 6.2.2), there were three distinct storms. Significant wave heights was calculated to be at 0.8m during these events, with a peak period of between 8 and 9 seconds. The time series of SPM concentrations indicates several scales of variability. In response to the spring/neap tidal cycle, background concentrations are higher during springs. The tidal response to SPM, comprises two components. The first is a semi-diurnal horizontal flux component arising from the advection of a horizontal turbidity gradient of fine material past the mooring site. Peaks in suspension arise when the tidal displacement has reached its maximum extent towards the north west at low water slack. Conversely minimum concentrations arise when this advective material reaches its maximum extent towards the south east on flood. This advective signal dominates the spring cycle and during storm conditions.

The second tidal response, seen clearly at neaps (run 120 - 170), is a quarter diurnal vertical flux component, driven by local resuspension of a more coarse material at maximum tidal streaming on both stages of the tide. The dip between the peaks is in phase with slack water, which suggests that material resuspended on flood is given time at slack water to settle out, only to be resuspended again on the reverse phase of the tide, giving rise to the second peak. Although, local resuspension is likely to occur during spring tides, its signal appears to have been swamped out by the stronger advective signal.

The presence of storm waves increases the amplitude of the tidally-varying concentrations, and will suspend both coarse and fine material into the water column. However whereas the coarse material settles out rapidly, there is a distinct time lag for the fine sediment

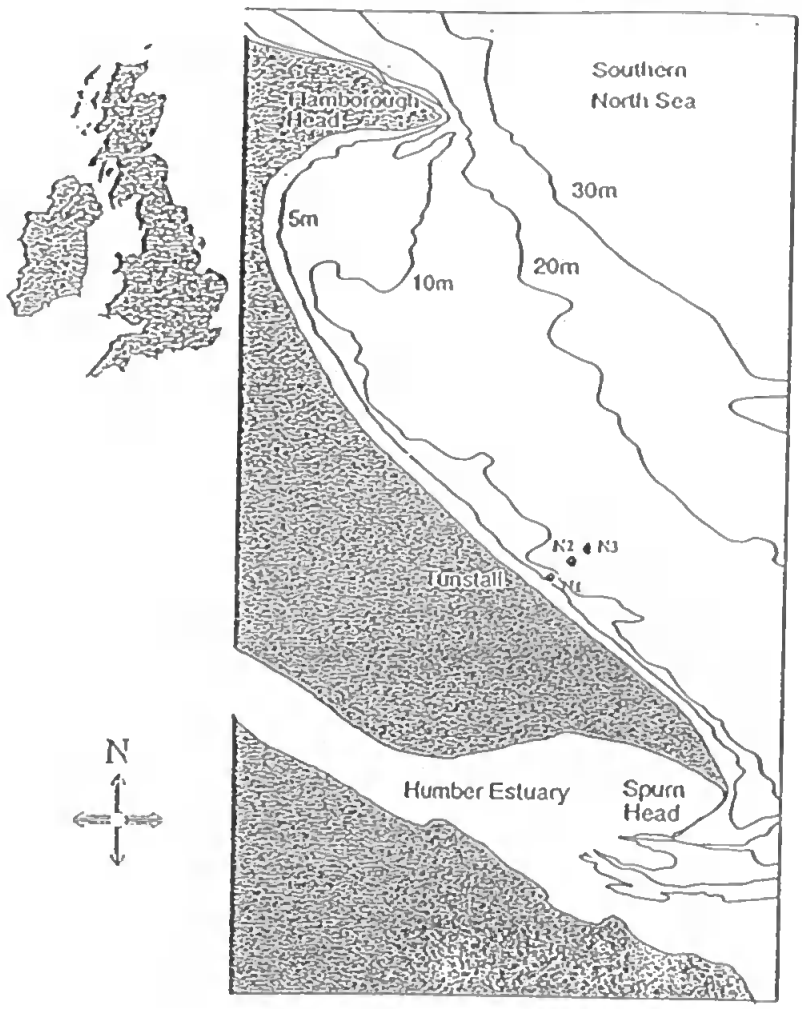


Figure 6.2.1. Holderness Coast

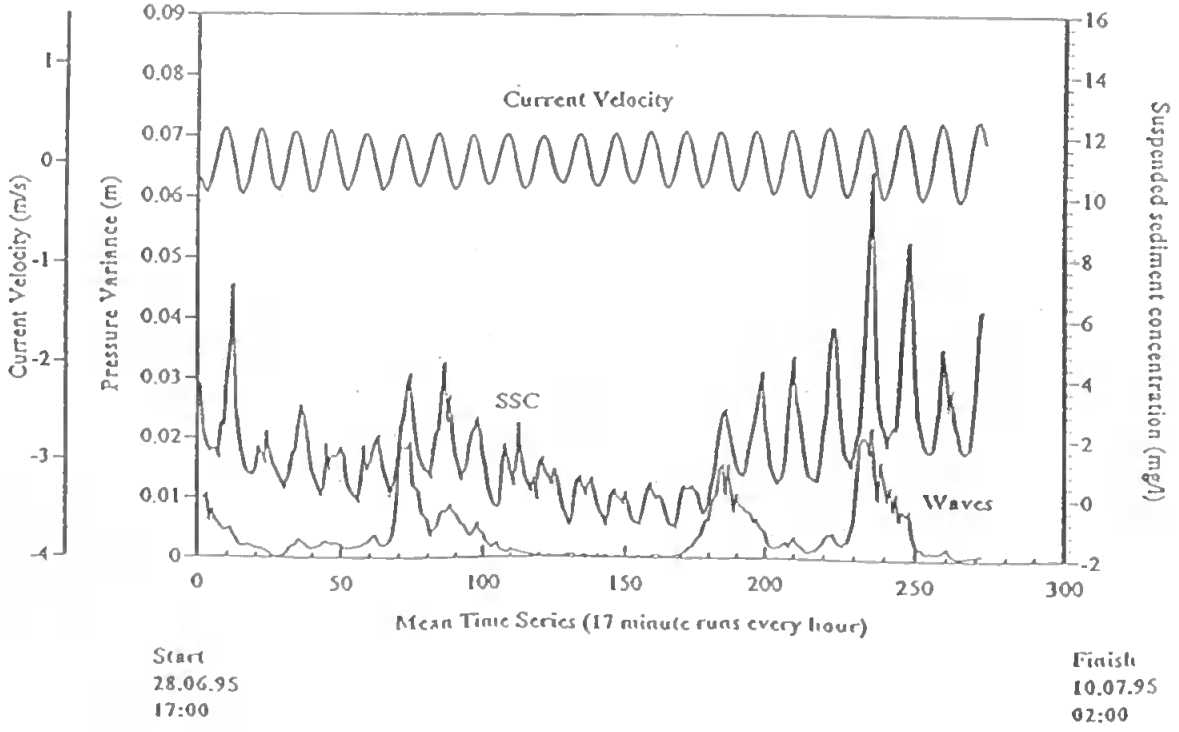


Figure 6.2.2. Burst Averaged Time Series of Current Velocity, Suspended Sediment Concentration (OBS) and Pressure Variance for July 1995—Inner Station.

fraction to settle out of suspension after the storm. It is evident from the time series that the concentration of this fine material takes a few days to reach pre storm background levels again. For any model to successfully predict these conditions off the Holderness Coast, it must not only take into account the different tidal processes, but it must also take into account this time lag memory aspect of the background sediment concentration in response to storms. (Chen, Blewett *et al.* 1997).

6.3 Current Velocity Model

In this section we focus on modelling the current velocity dynamics. The one reason we investigate the current velocity dynamics is that the most widely used sediment transport models are so-called cu-integral (concentration times velocity integral) type of models. The other reason is current velocity dynamics plays an very important part in sediment transport numerical model as we mentioned in Chapter 5. So current velocity is a very important variable to model when considering sediment transport. From Chapter 5, we know that the current velocity mainly is a function of its derivatives relavent to the direction (x, y, z) and some parameters, so in this section, the univariate model for current velocity profile is presented.

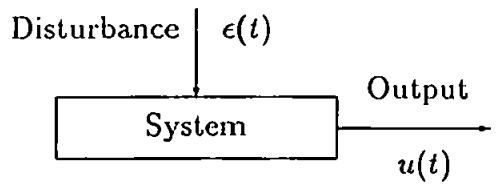


Fig 6.3.1 A Dynamic system with output $u(t)$ and disturbance $\epsilon(t)$ where t denotes time.

The distinguishing feature of a univariate time series current velocity model that no attempt is made to relate $u(t)$ to other variables except the uncontrollable disturbance $\epsilon(t)$. The variation in $u(t)$ assumed here is 'explained' solely in terms of its own past, although of course if $u(t)$ is dependent on space. The forecasts are then made by extrapolation. The statistical approach to forecasting is based on the construction of a model. The model defines a mechanism which is regarded as being capable of having produced the observations in question. Such a model of course when applied to the environment is

invariably stochastic.

We assume that the time series model here that describes the current velocity is a discrete one variable stochastic linear system and can be represented by the following AR(p) and ARMA(p, q) model.

6.3.1 AR Model

1. Model Description

First, we assume that the current velocity profile model is a one variable AR(p) model as follows:

$$A(z^{-1})u_n = w_n \tag{6.1}$$

where w_n is the system noise and the restriction on it are given in equations (3.66)–(3.67).

$$A(z^{-1}) = 1 + a_1 z^{-1} + ... + a_p z^{-p} \tag{6.2}$$

u_n is the current velocity at the time n, z^{-1} is a unit delay operator and $a_i, (i = 1, ..., p)$ are p unknown parameters to be estimated. Set

$$\theta^T = [-a_1, ..., -a_p] \tag{6.3}$$

$$x_n^T = [u_{n-1}, ..., u_{n-p}] \tag{6.4}$$

here θ_n is the estimate of θ at time n and it is easy to see that (6.1) also can be written by

$$u_n = \theta^T x_n + w_n \tag{6.5}$$

The recursive least squares algorithm of Chapter 3 (3.61–3.63) is used to identify the parameter vector θ.

2. Order determination

Definition 6.1.

σ^2 is defined as variance of one-step prediction error of time series model as follows:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} \tilde{u}_n^2 \qquad (6.6)$$

where $\tilde{u}_n = u_n - \hat{u}_n$, u_n is the real data and $\hat{u}_n = x_n^T \theta_{n-1}$ is the one-step prediction of the time series model at time n respectively.

Definition 6.2.

$$MPE = \max_{1 \le n \le N} |\tilde{u}_n| \qquad (6.7)$$

is defined the maximum one-step prediction error (MPE) of the time series model.

Definition 6.3.

$$MPV = \max_{\{1 \le n \le N\} \bigcup \{1 \le i \le d\}} |\tilde{\theta}_{i_n}| \qquad (6.8)$$

is defined the maximum parameter variation (MPV), where θ_{i_n} is the ith component of θ at time n.

The approach in our paper is to fit the model of progressively higher order, to calculate variance of one-step prediction error σ^2 for each value of order p, as well as to consider the MPE and MPV. The criterion is that if the addition of extra parameter matrices gives little improvement, we do not choose a higher order model.

The so-called F-test method (Söderström 1989) is used here to determine our model structure and the F-test results for our model candidates according to (4.11) and (4.12) are given as following in Table 6.3.1:

TABLE 6.3.1. The order comparison of one variable current velocity model

	MPE	MPV	σ^2
p=2	0.187765	0.050108	0.00100255
p=3	0.152344	0.0667977	0.000913569
p=4	0.153519	0.07542	0.000906986

From the data of Table 6.3.1, the model order selection procedures are as follows:

(i) Let AR(2)=$\mathcal{U}_1$, AR(3)=$\mathcal{U}_2$
$\chi^2_{0.05}(1) \approx 3.84$ and $x = 266 \times \frac{0.00100255-0.000913569}{0.000913569} = 25.9082 > 3.84$ reject AR(2)

(ii) Let AR(3)=$\mathcal{U}_1$, AR(4)=$\mathcal{U}_2$
$x = 266 \times \frac{0.000913569-0.000906986}{0.000906986} = 1.93066 < 3.84$ choose AR(3)

The AIC (4.13), FPE (4.14) and MERT (4.32) order determination test results are also given in the following Table 6.3.2.:

TABLE 6.3.2. The order determination of one variable current velocity model in different tests

Model Order	AIC	FPE	MERT
p=2	-1853.5	0.00101757	7.21259
p=3	-1869.5	0.000934254	152.907
p=4	-1862.44	0.00934575	2.333

It is shown in Table 6.3.2, that the AR(3) model is the best choice in all the three above mentioned order test methods in this case which is consistent with the F-test shown before. For simplicity, we just make use of F-test for our model order determination later in the thesis.

3. Simulation

According to the results of the model order determination, we choose following AR(3) model:

$$u_n = a_1 u_{n-1} + a_2 u_{n-2} + a_3 u_{n-3} + w_n \tag{6.9}$$

Set as:

$$\theta^T = [a_1, a_2, a_3] \tag{6.10}$$

$$x_n^T = [u_{n-1}, u_{n-2}, u_{n-3}] \tag{6.11}$$

For solving equation (6.9), the ELSM (3.61)-(3.63) is used (substitute u_n for y_n in (3.63)). From (6.2), $p = 3$. The time scale in here is one hour per run. Since the initial value is needed, the first three data as our initial value we really start the model at n=4 and we take $R_3 = 3I_3$. The computation procedure of ELSM is as follows:

(i) Construct x_n according to (6.11).
(ii) Select initial values of θ_3 and R_3.
(iii) Calculate K_n, R_n and θ_n according to (3.61)-(3.63) based on the K_{n-1}, R_{n-1} , θ_{n-1} and x_n ($n \geq 3$).

The simulation results are given as following:

(i) **TABLE 6.3.3. The Parameter Estimation of AR(3) Model**

Parameter	Mean	Standard Deviation
a_1	1.37709	0.01222716
a_2	-0.39324	0.00640278
a_3	-0.328613	0.00816693

(ii) The comparison of model with the data set is given in Figure 6.3.2 which shows the AR(3) model presented here has a good description of the model and data fitting.

(iii) The parameters variation is shown in Figure 6.3.3. From Figure 6.3.3 and Table 6.3.3, all three parameters of AR(3) can be considered as time-invariant parameter and have good convergence properties which means the model structure we chose is quite reseanable to describe the current velocity dynamics.

(iv) The model prediction or forecasting technique which we mentioned in subsection 2.3.3 is used here for long term prediction of AR(3) current velocity dynamic model. Under the assumption that we don't know the current velocity data $u_n, n \geq 223$, and let the white system noise series w_n meet the conditions $Ew_n = 0, Ew_n^2 = 0.00083$. We give the optimum k-step prediction $f_{222,k}, k = 1, 2, ..., 48$ based on the current velocity data

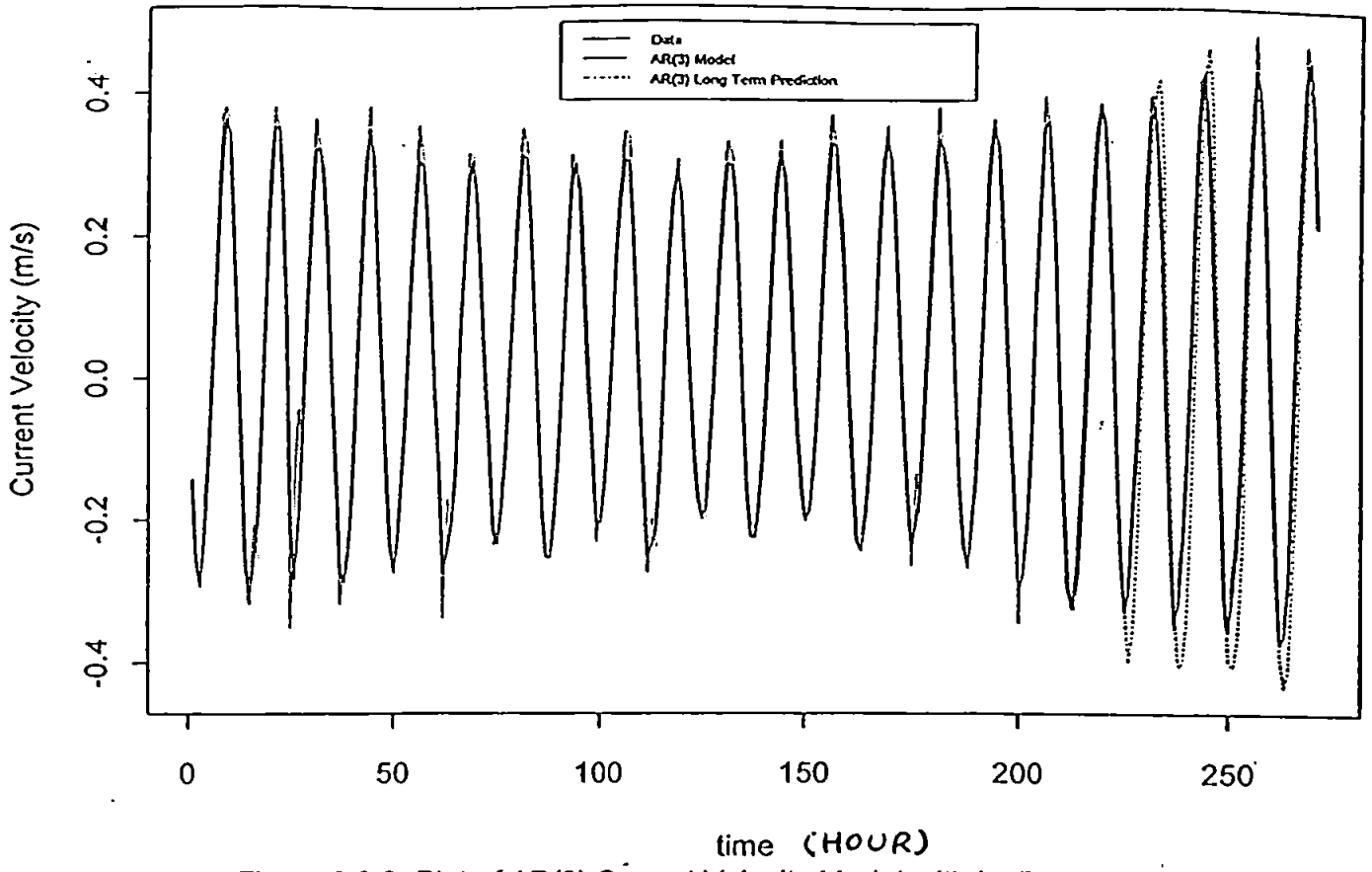


Figure 6.3.2. Plot of AR(3) Current Velocity Model with its Prediction vs Data

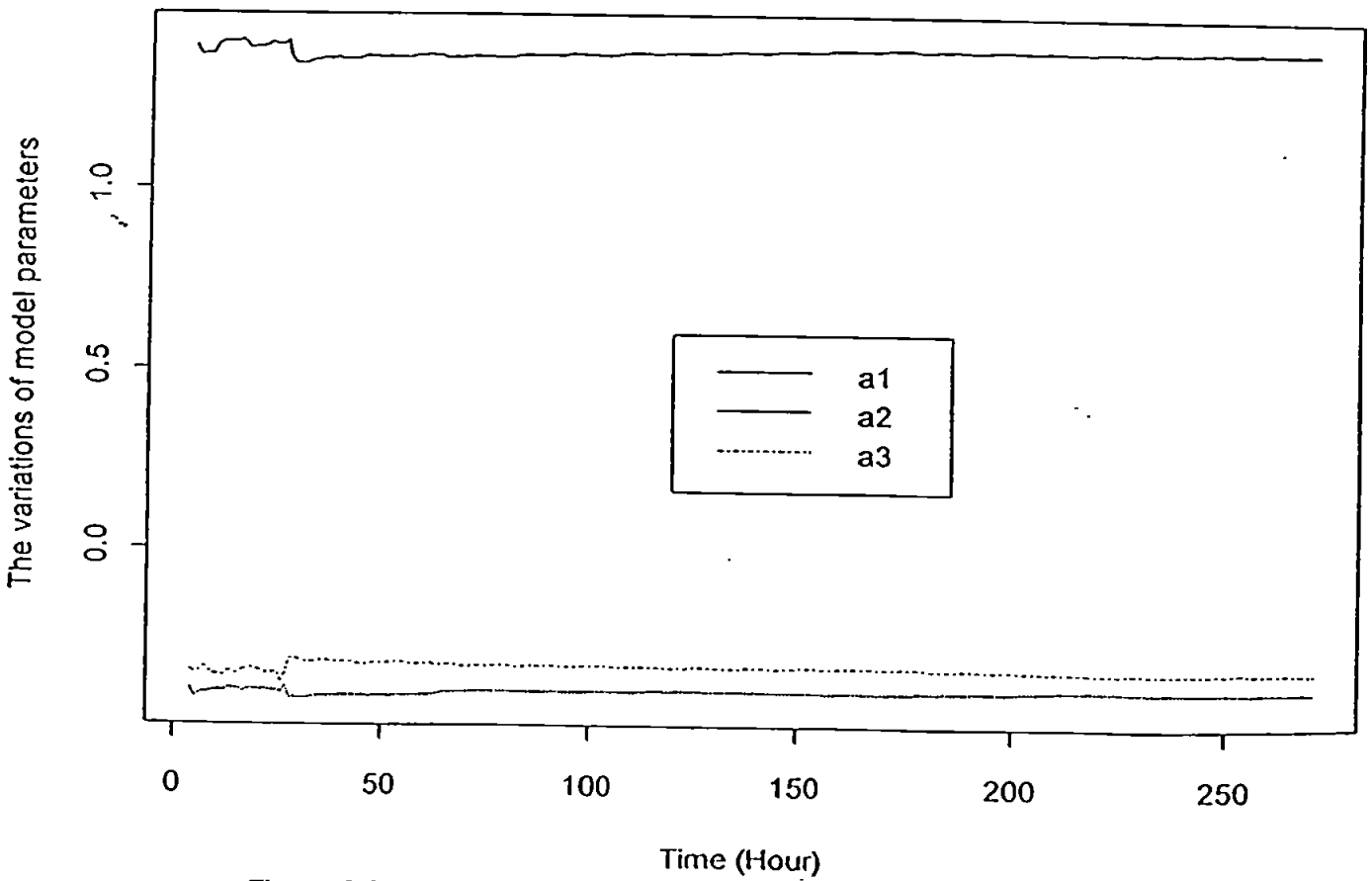


Figure 6.3.3. Plot of parameters variation of AR(3) Current Velocity Model

$u_n, n \leq 222$ and w_n. The model prediction results given in the Figure 6.3.3. which shows the AR(3) model presented here coincides with both the periodical and trend appearance of the current velocity data set very well.

6.3.2 ARMA Model

1. Model Description

In last subsection, we set up a AR(p) model for current velocity dynamics which assumed the current velocity dynamic system is disturbed by white noise series (*i.e.* $\epsilon_n = w_n$). In this subsecton, we try to consider more general kind of model (*i.e.* ARMA(p, q) model) for the current velocity dynamics and assume that the current velocity dynamic system is disturbed by the colored noise (*i.e* $\epsilon_n = w_n + c_1 w_{n-1}$) instead of white noise in AR(p) model. For simplicity, we set $q = 1$ and ARMA($p, 1$) model is represented as following:

$$A(z^{-1})u_{n+1} = w_{n+1} + c_1 w_n \tag{6.12}$$

where u_n is the current velocity, w_n is the system noise and the restriction on it are given in equations (3.66)–(3.67).

$$A(z^{-1}) = 1 + a_1 z^{-1} + ... + a_p z^{-p} \tag{6.13}$$

z^{-1} is a unit delay operator, $a_i, (i = 1, ..., p)$ and c_1 are $p + 1$ unknown parameters to be estimated. Set

$$\theta^T = [-a_1, ..., -a_p, c_1] \tag{6.14}$$

$$x_n^T = [u_{n-1}, ..., u_{n-p}, e_{n-1}] \tag{6.15}$$

here θ_n is the estimate of θ at time n, $e_n = u_n - \theta_n^T x_n$ as the estimation of w_n and it is easy to see that (6.12) also can be written by

$$u_n = \theta^T x_n + w_n + c_1(w_{n-1} - e_{n-1}) \tag{6.16}$$

The recursive least squares algorithm of Chapter 3 (3.55–3.57) is used to identify the parameter vector θ.

2. Order determination

The order determimation here is similar to the last subsection *i.e.* to fit the model of progressively higher order, to calculate variance of one-step prediction error σ^2 for each value of order p, as well as to consider the MPE and MPV. The criterion is that if the addition of extra parameter matrices gives little improvement, we do not choose a higher order model.
The F-test results for our model candidates according to (4.11) and (4.12) are given as following in Table 6.3.4:

TABLE 6.3.4. The order comparison of ARMA(p,1) current velocity model

	MPE	MPV	σ^2
p=1	0.283617	0.121979	0.0067109
p=2	0.145547	0.0669972	0.000898962
p=3	0.152792	0.0787776	0.000919854

(i) Let ARMA(1,1)=$\mathcal{U}_1$, ARMR(2,1)=$\mathcal{U}_2$
$\chi^2_{0.05}(1) \approx 3.84$ and $x = 266 \times \frac{0.0067109-0.000898962}{0.000898962} = 1719.73 > 3.84$ reject ARMA(1,1).

(ii) Let ARMA(2,1)=$\mathcal{U}_1$, ARMA(3,1)=$\mathcal{U}_2$
$x = 266 \times \frac{0.000898962-0.000919854}{0.000919854} = -6.041472 < 3.84$ choose ARMA(2,1).

3. Simulation

In here, we choose which is a simpler form model of (6.12) as following ARMA(2,1) model from our model order selection:

$$u_{n+1} = a_1 u_n + a_2 u_{n-1} + w_{n+1} + c_1 w_n \tag{6.17}$$

Set as:

$$\theta^T = [a_1, a_2, c_1] \tag{6.18}$$

$$x_n^T = [u_{n-1}, u_{n-2}, e_{n-1}] \tag{6.19}$$

To solve equation (6.17), the ELSM (3.61)-(3.64) is applied (substitute u_n for y_n) in (3.63) and (3.64)). From (6.17), $p = 2$ and the time scale in here is one hour per run. Since the initial value are need, the first two data as our initial value we really start the model at n=3.

The computation procedure of ELSM is as follows:

(i) Construct x_n according to (6.19) and (3.64), ($n \geq 2$).
(ii) Select initial values of θ_2 and R_2.
(iii) Calculate K_n, R_n and θ_n according to (3.61)-(3.63) based on the K_{n-1}, R_{n-1} , θ_{n-1} and x_n ($n \geq 2$).

The simulation results are given as follows:

(i) **TABLE 6.3.5. The Parameter Estimation of ARMA(2,1) Model**

Parameter	Mean	Standard Deviation
a_1	1.7239	0.0128287
a_2	-0.977191	0.0075229
c_1	-0.383821	0.00781896

(ii) The comparison of model with data set is given in in Figure 6.3.4. which shows the ARMA(2,1) model presented here is reasonably good both in data fitting and system description.

(iii) The parameters variation is shown in Figure 6.3.5. All the three parameters in ARMA(2,1) model have good convergence and time-invariant properties which means the model we choosed here is a good one to describe the current velocity dynamics near the Holderness Coast, England.

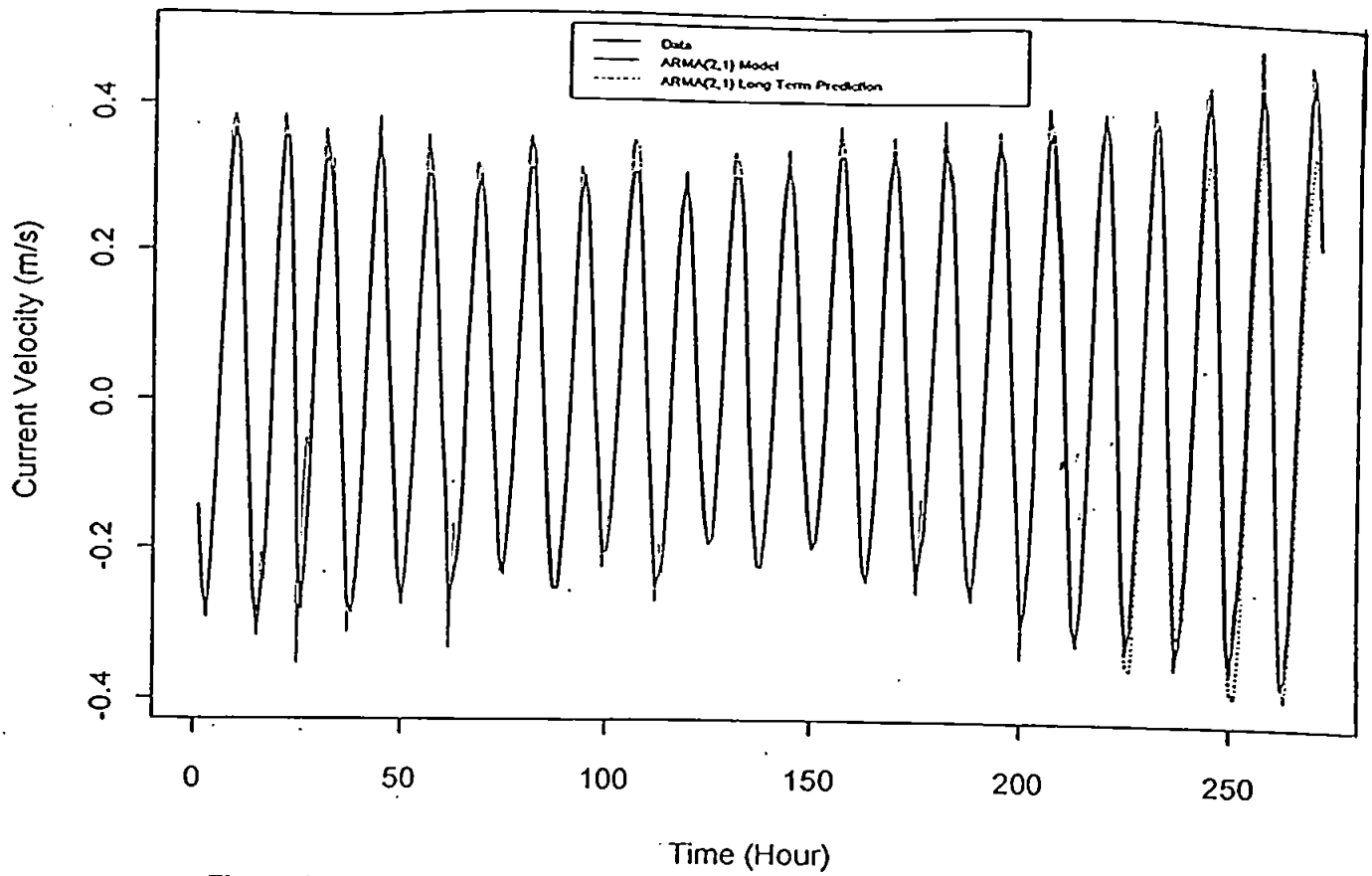


Figure 6.3.4. Plot of ARMA(2,1) Current Velocity Model with its Prediction vs Data

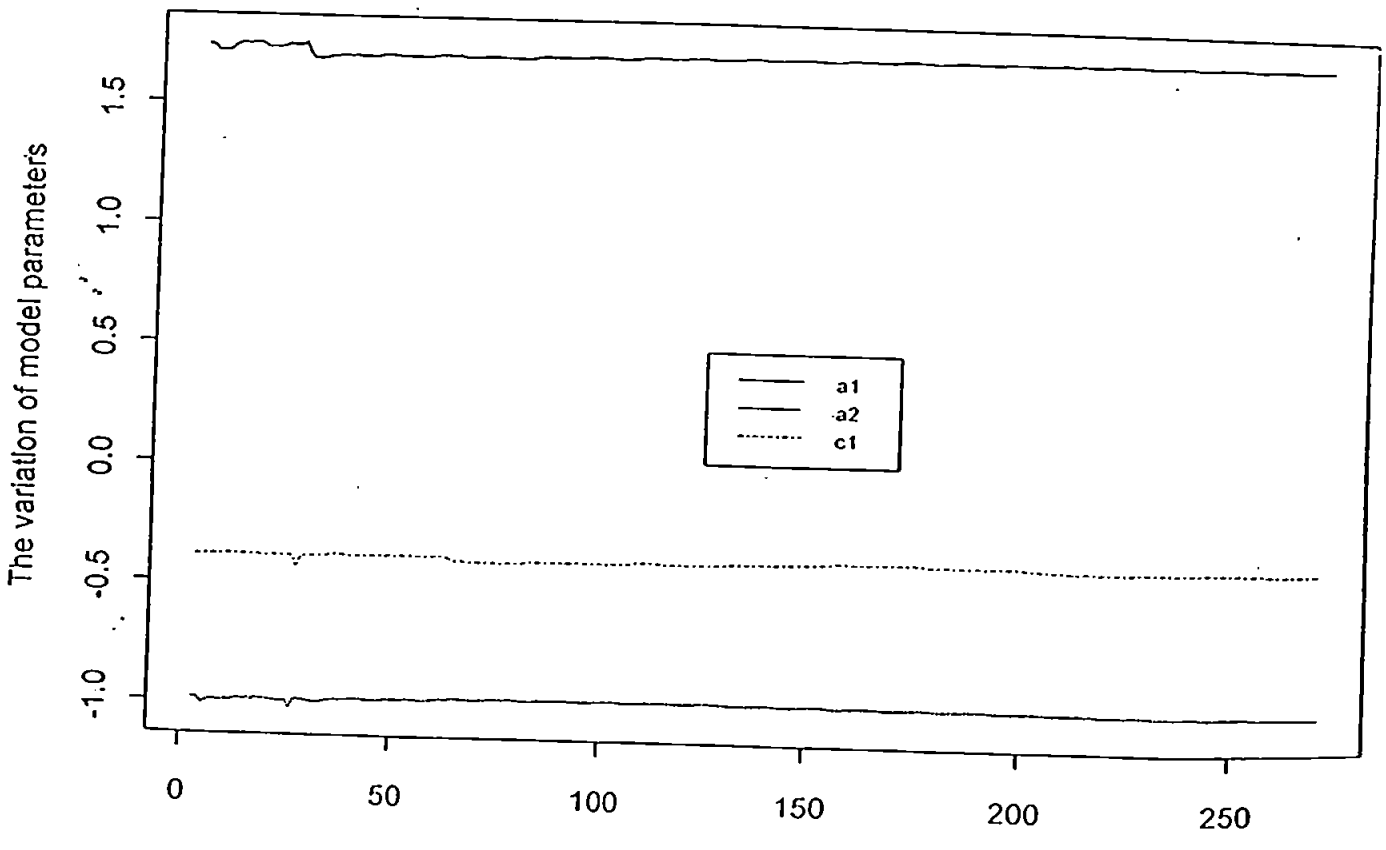


Figure 6.3.5. Plot of parameters variation of ARMA(2,1) Current Velocity Model

(iv) The model peridiction method given in subsection 2.3.4. is used here for long term prediction of ARMA (2,1) current velocity dynamic model. Under the assumption that we don't know the current velocity data $u_n, n \geq 224$, and the noise series w_n is the same as that we used in last subsection long term prediction in AR(3) model. We give the optimum k-step prediction $f_{223,k}, k = 1, 2, ..., 48$ based on the current velocity data $u_n, n \leq 223$. The model prediction results are also given in the Figure 6.3.5, which shows the ARMA(2,1) model presented here works very well in 48 hours prediction compared with data set.

6.3.3 Comparison of AR(3) Model with ARMA(2,1) Current Velocity Models

In last two subsections, we present two kinds of time series current velocity model which both work very well as shown in the simulations. The comparison of model fitting, parameter variation and long term prediction of AR(3) and ARMA(2,1) are discussed in this subsection. In order to compare the different kinds of models, the following definitions are given.

Definition 6.4

The mean of long term prediction error (MLPE), the standard deviation of long term prediction error (SDLPE) and the maximum prediction error (MPR) are defined respectively as follows:

$$MLPE = \frac{1}{48}\sum_{k=1}^{48}|f_{n,k} - u_{n+k}| \tag{6.20}$$

$$SDLPE = \sqrt{\frac{1}{48}\sum_{k=1}^{48}(f_{n,k} - u_{n+k})^2} \tag{6.21}$$

$$MPR = \max_{1\leq k\leq 48}|f_{n,k} - u_{n+k}| \tag{6.22}$$

The simulation results of the two kind of time series current velocity models are presented in Table 6.3.6.

TABLE 6.3.6. Comparison of Current Velocity Models

	MPE	MPV	σ^2	MLPE	SDLPE	MPR
AR(3)	0.152344	0.0667977	0.000913569	0.0940563	0.0597524	0.220193
ARMA(2,1)	0.145547	0.0669972	0.000898962	0.051567	0.0368606	0.138268

From Table 6.3.6, all the items of ARMA(2,1) are smaller than those of AR(3) except that the MPV of ARMA(2,1) is nearly same as the one of AR(3). It is shown that ARMA(2,1) has better simulation results than AR(3) especially in the case of long term prediction. This is because the ARMA(2,1) takes advantage of the more imformation from the model error. It is also shown that the consideration of the coloured system noise has improved the model on the whole and suggested the ARMA(2,1) is a more reasonable choice in this special case.

6.4 Suspended Sediment Concentration Model

6.4.1 Introduction

A series of models has been produced to simulate these processes using time series observations. For processes in the Southern North Sea, Jago and Jones (1993) formulated a conceptual model, which combined resuspension and advection components superimposed on a background concentration. The model assumed horizontal homogeneity for the resuspension component and a negligible settling rate for the background component. Hence, the resuspension component is a simple function of current speed, and the advection component is a function of tidal displacement. The model however, provided little insight into the physical processes of SPM resuspension. It simulated conditions most successfully in the upper water layers, but because the resuspension criterion depended on current volocity and without a threshold shear stress, it was not entirely successful in its representation of particle entrainment from the seabed. However, this limitation became less serious higher up in the water column, where vertical diffusion, rather than bed erosion rate became dominant. What became clear from these results is that a more physically orientated and complex model is required to accurately simulate the near-bed region.

Further investigations for the Southern North Sea led Jago *et al.* (1993) to produce of a one-dimensional turbulence model for particle resuspension and flux which incorporated the effects of both horizontal advection of a prescribed concentration and vertical diffusion. Their model is based on a hydrodynamic model presented by Simpson and Sharples (1991) and simulations were run with both an unlimited and a limited supply of resuspendable material assuming that all the available material is already in suspension at the start of the model run (at maximum tidal streaming), produced results which agreed with observation in both phase and form (Jago *et al.* 1993). This type of model however, could not predict horizontal concentration gradients of the finer material and would need to include tidal advection of prescribed horizontal gradients (Jones *et al.* 1994).

What is ignored in both these models, is the influence of wave activity and storm events on the resuspension processes. Recent work by Green *et al.* (1990) simulated enhanced bed shear stress in a wave/current flow with observations from a coastal site in the Southern North Sea, and showed that the peak wave stress may be significantly increased by non-linear interaction with the tidal current even under small waves, and increased resuspension of bed material due to wave enhancement of bed stress was apparent. Observations taken off the Holderness Coast in the Southern North Sea, show that storm activity acts to enhance the background levels of fine material in the water column. Whereas the coarse material settles out rapidly, the fine material exhibits a distinct time lag, persisting in the water column days after the storm has passed.

In this section, the suspended sediment dynamics is taken as a unknown stochastic system and a new model is developed to simulate both horizontal advection, local resuspension and the effects of waves and storms. The model considers and includes the following as stochastic processes; current velocity, wave variation, tidal displacement and suspended sediment concentration. The first three are set as inputs to the system model, and the last one is considered as the output of the system model. System Identification theory is applied in the model to identify the unknown parameters of the model, based on in situ data collected in the study area. In this case, the system identification technique is applied to the Holderness field data collected in July, 1995 by Blewett to illustrate the relationship between suspended sediment concentration, current velocity, tidal displacement and varying wave conditions. The aim is to show that this particular theory is suitable for matching known data sets, and this will be achieved by ensuring all the parameters remain virtually constant when tackling future data sets. By operating the model in this way, the results obtained should simulate the data with some degree of accuracy.

Chen and Dyke (1995) have already presented a single variable time series model (ARMA model) for the current velocity and suspended particulate matter (SPM) concentration for the Tamar Estuary and it shows very good agreement with the data collected from the Tamar Estuary during the high water slack by Fennessy *et al.* (1994). In Chen and Dyke (1996), a multivariable stochastic time series model (multivariable ARMA Model) is set up to describe the suspended sediment concentration and current velocity over a depth profile respectively. The recursive least squares identification algorithm is used to identify the unknown parameter matrices of the model and the simulations are given to show the a good approximation to real data collected from the Rufiji Delta, Tanzania by Fisher (1994).

In the models of Chen and Dyke (1995) and Chen and Dyke (1996) the suspended sediment concentration variation is assumed to depend on its own past values and uncontrollable system noises. By contrast, in here, the suspended sediment concentration is assumed to be related to the current velocity, tidal displacement and wave variation in addition to its own past values, in order to produce a more physically based model. These model results are compared with those from a conventional multiple regression model using the same external input variables but not including past SPM values. The simulation results show that the new model developed produces good agreement with the real data collected from the Holderness Coast, England, whilst the simple regression model generally gives poor agreement.

6.4.2 Multiple Input Single Output (MISO) Models

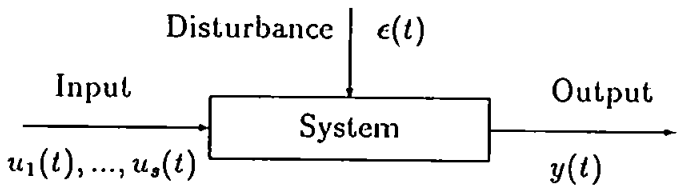


Fig 6.4.1. A Dynamic system with inputs $u_1(t), ..., u_s(t)$, output $y(t)$ and disturbance $\epsilon(t)$ where t denotes time.

Figure 6.4.1 shows a schematic of the type of predictive model under discussion. Multiple input time series, $u_1(t), ..., u_s(t)$, produce a single output time series, $y(t)$ through a system box which receives disturbance (noise), $\epsilon(t)$, so that the relationship between input and

output is stochastic. The characteristics of the system are assumed to be time-invariant and linear, and the system output can depend upon earlier values of both input and output time series. Such a system is known as a multi-input single-output (MISO) model. In the present case, the output time series is the suspended sediment concentration measured by the Boundary Layer Intelligent Sensor System (BLISS) tripods, and the input time series are the current velocity, the tidal displacement of water along the coast (a time integral of the alongcoast velocity), and the wave intensity (each of which is shown in Figure 6.2.2).

For s different input time series, a MISO system can be written mathematically in the form:

$$\begin{aligned} y_n + a_1 y_{n-1} + ... + a_p y_{n-p} \;=\; & b_{1_1} u_{1_{n-1}} + ... + b_{1_{n-q_1}} u_{1_{n-q_1}} \\ & + b_{2_1} u_{2_{n-1}} + ... + b_{2_{n-q_2}} u_{2_{n-q_2}} \\ & ..... \\ & + b_{s_1} u_{s_{n-1}} + ... + b_{s_{n-q_s}} u_{s_{n-q_s}} \\ & + w_n + c_1 w_{n-1} + ... + c_r w_{n-r} \end{aligned} \tag{6.23}$$

y_n and $u_{i_n} (n = 1, 2, 3, ...)$ are the output and i-th input of the system respectively. p is the order of the system and $w_n, (n = 1, 2, 3, ...)$ is a noise series and a_i, b_{j_k} and $c_l (i = 1, ..., p; j = 1, ...s; k = 1, ..., q_j; (q_j \leq p); l = 1, ...r.)$ are unknown parameters to be estimated.

Let z^{-1} is a unit delay operator and

$$A(z^{-1}) = 1 + a_1 z^{-1} + ... + a_p z^{-p} \tag{6.24}$$

$$B_i(z^{-1}) = b_{i_1} z^{-1} + ... + b_{i_{q_i}} z^{-q_i} \tag{6.25}$$

$$C(z^{-1}) = 1 + c_1 z^{-1} + ... + c_r z^{-r}. \tag{6.26}$$

The equation (6.23) can be written as follows:

$$A(z^{-1}) y_n = \sum_{i=1}^{s} B_i(z^{-1}) u_{i_n} + \epsilon_n \tag{6.27}$$

where $\epsilon_n = C(z^{-1}) w_n$ is the system noise and set

$$\theta^T = [-a_1, ..., -a_p, b_{1_1}, ..., b_{1_{q_1}}, b_{2_1}, ..., b_{s_{q_s}}, c_1, ..., c_r] \tag{6.28}$$

$$x_n^T = [y_{n-1}, ..., y_{n-p}, u_{1_n}, ..., u_{1_{n-q_1}}, u_{2_n}, ..., u_{s_{n-q_s}}, e_{n-1}, ..., e_{n-r}] \tag{6.29}$$

$$e_n = y_n - \theta_n^T x_n \tag{6.30}$$

$$d = p + r + \sum_{i=1}^{s} q_i \tag{6.31}$$

here θ is the true parameter matrix, x_n is the regression vector consisted of the information of input, output and the estimation of system noise, e_n is the estimation of $\dot{w}_n$ and θ_n is the estimate of θ at time n . It is easy to see that (6.23) or (6.27) also can be written as

$$y_n = \theta^T x_n + C(z^{-1})w_n + e_n - C(z^{-1})e_n. \tag{6.32}$$

The first term of the right side of equation (6.32) can be considered as the estimation of y_n (since we do not know the true θ, y_n is estimated by $\theta_n^T x_n$) and the remaining terms on the right hand side can be considered as a kind of filter of system noise.

6.4.3 Order Determination

For simplicity, we choose a MISO time series model which is a simpler form of (6.27), using the following MISO(p) model for whole process:

$$y_n = a_1 y_{n-1} + \ldots + a_p y_{n-p} + b u_{n-1}^2 + c T_{n-1} + f T_{n-1} W_{n-1} + w_n. \tag{6.33}$$

The appropriate form of equation (6.28) and (6.29) then becomes:

$$\theta^T = [a_1, \ldots, a_p, b, c, f] \tag{6.34}$$

$$x_n^T = [y_{n-1}, \ldots, y_{n-p}, u_{n-1}^2, T_{n-1}, T_{n-1} W_{n-1}] \tag{6.35}$$

where, y_n is suspended sediment concentration (m/l), u_n is the current velocity (m/s), T_n is the along-coast tidal displacement (km) and W_n is the wave elevation variance at time n respectively. In equation (6.33), the resuspension signal is modelled by the term in u^2, and the tidal advection effect is modelled by two terms, one simply proportional to the tidal displacement and a second term multiplying the displacement by the wave variance, to simulate the increase in the advection signal during storms.

To solve equation (6.33), the recursive least squares algorithm (3.61-3.63) is used. From (6.31), where $q_i = 1, i = 1, 2, 3.$, r=0 and $d = p + 3$ and the time scale here is one hour per run. Since the initial values are required, we used the first p data as our initial values and therefore start the model process at $n = p + 1$.

The approach in our paper is to fit models of progressively higher order, to calculate variance of one-step prediction error (the residual sum of squares (RSS) in traditional linear regression model) for each value of order p, as well as to consider the MPE and MPV. The criterion is that if the addition of extra parameter matrices gives little improvement, we do not choose a higher order model.

The F-test results for our model candidates according to (4.11)–(4.12) are given in Tables 6.4.1–6.4.3:

TABLE 6.4.1. The Order Comparison in Whole Period ($0 \leq n \leq 271$)

Order	MPE	σ^2	MPV
p=1	4.36456	0.573506	2.59355
p=2	3.87259	0.485172	2.84562
p=3	5.42097	0.518003	2.9651

Let MISO(1)=$\mathcal{U}_1$, MISO(2)=$\mathcal{U}_2$

$\chi^2_{0.05}(1) \approx 3.84$ and $x = 268 \times \frac{0.573506-0.485172}{0.485172} = 49.7941$ reject MISO(1).

Let MISO(2)=$\mathcal{U}_1$, MISO(3)=$\mathcal{U}_2$

$\chi^2_{0.05}(1) \approx 3.84$ and $x = 268 \times \frac{0.485172-0.518003}{0.518003} = -16.9858$ choose MISO(2).

TABLE 6.4.2. The Order Comparison in Calm Period ($110 \leq n \leq 170$)

Order	MPE	σ^2	MPV
p=1	1.072314	0.1268888	0.947681
p=2	0.771115	0.068666	0.435191
p=3	0.596783	0.0649163	0.341153

Let MISO(1)=$\mathcal{U}_1$, MISO(2)=$\mathcal{U}_2$

$\chi^2_{0.05}(1) \approx 3.84$ and $x = 57 \times \frac{0.1268888-0.068666}{0.068666} = 49.179$ reject MISO(1).

Let MISO(2)=$\mathcal{U}_1$, MISO(3)=$\mathcal{U}_2$

$\chi^2_{0.05}(1) \approx 3.84$ and $x = 57 \times \frac{0.068666-0.0649163}{0.0649163} = 3.29244$ choose MISO(2).

TABLE 6.4.3. The Order Comparison in Storm Period ($171 \leq n \leq 270$)

Order	MPE	σ^2	MPV
p=1	3.88889	2.77873	2.91095
p=2	2.97176	0.523499	1.2119
p=3	4.73119	2.84352	2.89949

Let MISO(1)=$\mathcal{U}_1$, MISO(2)=$\mathcal{U}_2$
$\chi^2_{0.05}(1) \approx 3.84$ and $x = 97 \times \frac{2.77873-0.523499}{0.523499} = 417.876$ reject MISO(1).

Let MISO(2)=$\mathcal{U}_1$, MISO(3)=$\mathcal{U}_2$
$\chi^2_{0.05}(1) \approx 3.84$ and $x = 97 \times \frac{0.523499-2.84352}{2.84352} = -79.1421$ choose MISO(2).

Thus, for all three time period a second order model is adequate at the 95% level.

6.4.4 Parameter Estimation

From order determination we choose MISO(2) time series model for all three time periods in the following form:

$$y_n = a_1 y_{n-1} + a_2 y_{n-2} + b u^2_{n-1} + c T_{n-1} + f T_{n-1} W_{n-1} + w_n \tag{6.36}$$

with

$$\theta^T = [a_1, a_2, b, c, f] \tag{6.37}$$

$$x_n^T = [y_{n-1}, y_{n-2}, u^2_{n-1}, T_{n-1}, T_{n-1} W_{n-1}] \tag{6.38}$$

Equation (6.36) is using the Recursive Least Square Method (3.61-3.63), where for order 2, d=5, and the model simulation begin at n=3.

In the recursive (also called on-line) identification method we used here, the parameter estimates are computed recursively in time similar to the ELSM presented in section 6.3. and the parameter estimation results are given as following:

TABLE 6.4.4 The Parameter Estimation in Whole Period

Parameter	Mean	Standard Deviation
a_1	0.915355	0.108976
a_2	0.0310859	0.11581
b	0.377286	0.2565
c	6.55802	0.405544
f	0.0468373	0.0415331

TABLE 6.4.5. The Parameter Estimation in Calm Period

Parameter	Mean	Standard Deviation
a_1	0.873825	0.0481795
a_2	0.107592	0.0503243
b	-0.0780674	0.0950974
c	1.96009	0.0643295
f	0.674854	0.110728

TABLE 6.4.6. The Parameter Estimation in Storm Period

Parameter	Mean	Standard Deviation
a_1	1.23308	0.117625
a_2	-0.278993	0.116848
b	0.259172	0.122158
c	8.17912	0.304354
f	-0.00410298	0.0293338

Figures 6.4.2–6.4.4 show the comparison of the MISO(2) suspended sediment concentration models and data for the three time periods.

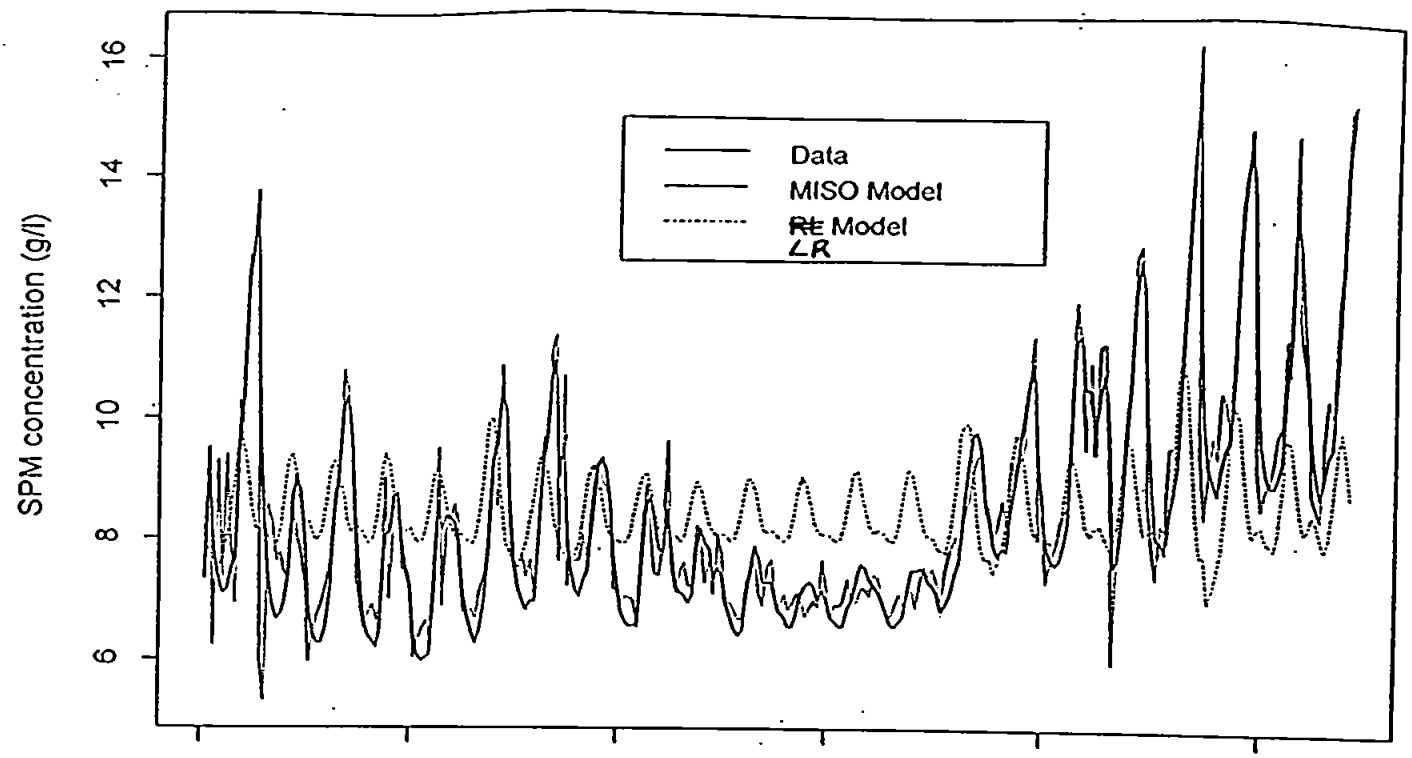


Figure 6.4.2. Plot of Models vs Data (Whole Period)

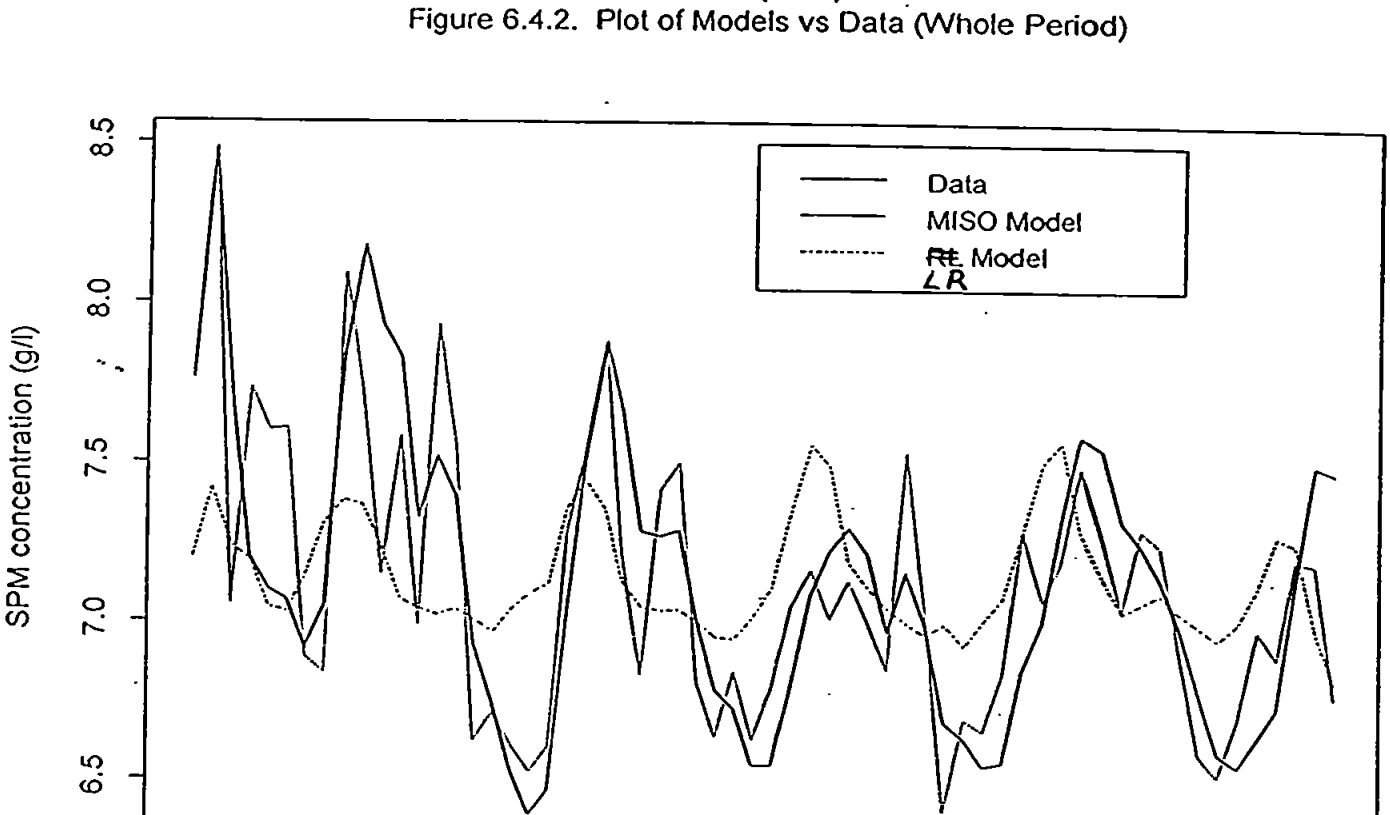


Figure 6.4.3. Plot of Models vs Data (Calm Period)

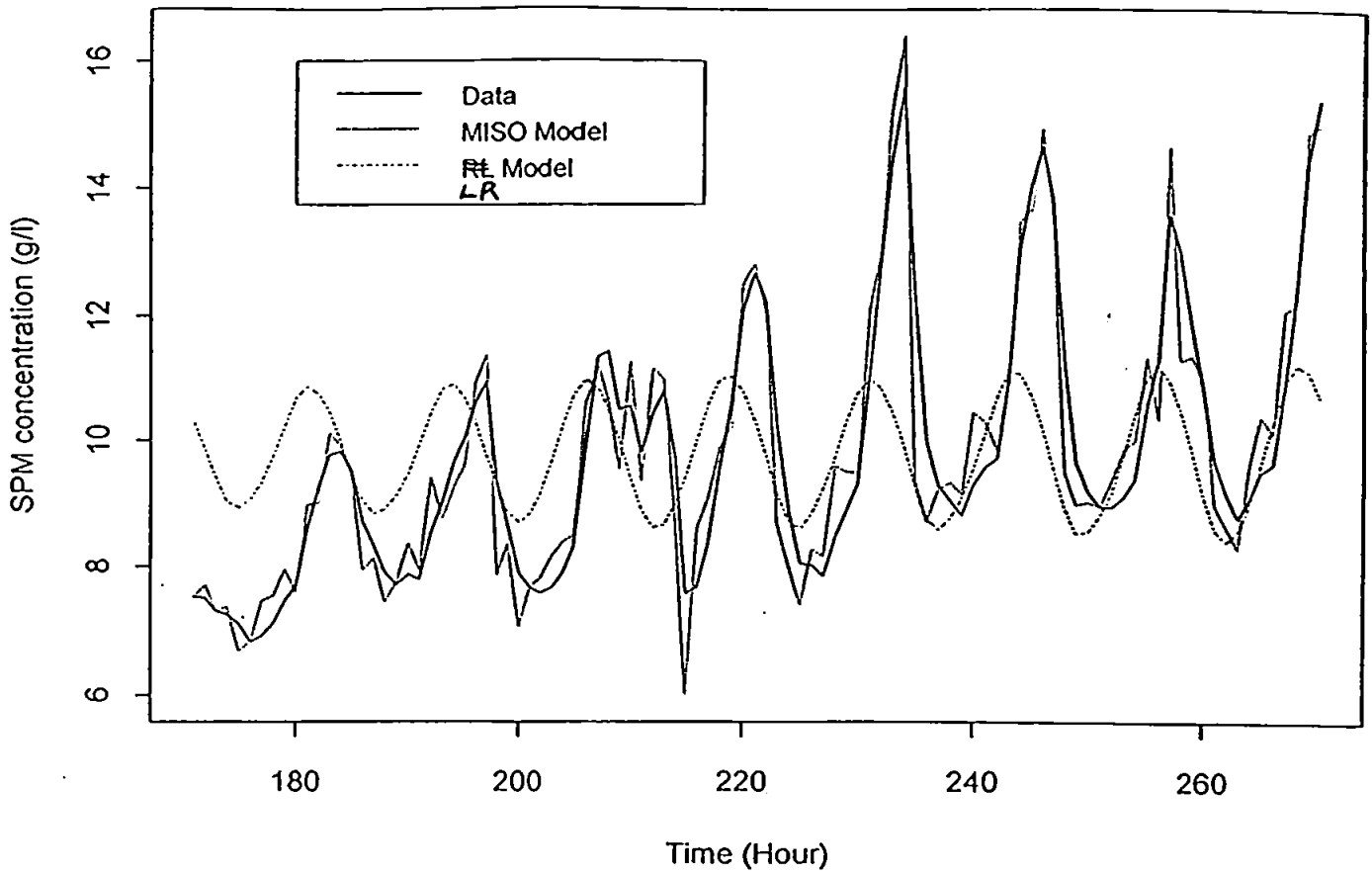


Figure 6.4.4. Plot of Models vs Data (Storm Period)

6.4.5 Comparison with Linear Regression Model (LR)

The traditional linear regression model is as follows:

$$y_n = a + bT_n + cu_n^2 + fW_n * T_n \tag{6.39}$$

The fitted parameters with their standard deviation (shown in brackets) and the model comparison of all time periods are given in Table 6.4.7. and Table 6.4.8. respectively.

TABLE 6.4.7. The Fitted Parameters In the Regression Model

Model	a	b	c	f
Whole Process	8.08 (0.17)	0.34 (0.19)	6.02 (2.73)	66.21 (29.32)
Calm Process	7.07 (0.09)	0.71 (0.26)	1.90 (2.04)	-1148.5 (459.63)
Storm Process	9.66 (0.30)	0.92 (0.31)	0.43 (3.9)	3.61 (39.24)

TABLE 6.4.8. The Model Comparison

Model	MISO(Whole)	LR(Whole)	MISO(Calm)	LR(Calm)	MISO(Storm)	LR(Storm)
σ^2 (RSS)	0.4852	3.1916	0.0687	0.1532	0.5235	3.4779

From Tables 6.4.7–6.4.8. and Figures 6.4.2–6.4.4, it is shown that the MISO time series model is much better than the linear regression model in data fitting and prediction. This is because the MISO model takes advantage of the on-line identification method, allowing variation of suspended sediment concentration to depend on past values as well as concurrent hydrodynamic conditions.

6.5 Discussion and Conclusion

Overall the single variable AR, ARMA current velocity model presented here is seen to fit the measured current velocity data off the Holderness coast very well. They provide a

novel modelling technique in current velocity. Since the recursive least square identification method is used, the model can be instantaneous response to change in hydrodynamic conditions and adaptive to the system variation. Thus the improvement of the model and model prediction are made. From the simulation results presented in Figures 6.3.2–6.3.5, the parameters of the model have very good convergency in the whole period, which means both the model structure and description assumed here are very reasonable and the model has sound reliability of long term prediction.

In the modelling of SPM concentration, the MISO model is also seen to fit the measured variation of suspended sediment concentration off the Holderness coast very well. It provides a major improvement on the more traditional linear regression (LR) model, because it relaxes the LR requirement of instantaneous response to changing hydrodynamic conditions. The improvement occurs for all time scales. MISO accurately follows the very significant tidal-averaged background concentration changes which appear to depend on both storm conditions and the spring-neap cycle: LR shows much smaller changes than observed. MISO gives good agreement with the changing amplitude of the semi-diurnal signal, though tending to overshoot at the extreme values, particularly minimum concentrations; LR shows some of the observed variation, but at a lower level than observed. Finally MISO gets much closer to the observed quarter-diurnal concentration variation during the calm period, attributed to local resuspension. There is some evidence that the MISO model produces some spurious high frequency variation not seen in the observations; time period 135–180 in Figure 6.4.2. is a particularly clear example, as is the tendency to overshoot the extremes. However the dominant features of the time series are vastly better predicted by the MISO model when compared to the LR model.
One drawback of the MISO method is that, as a model-fitting exercise, it is only strictly applicable to the condition fitted. It is clear from the results that the best-fit parameters differ significantly between the storm and the calm periods, and that the use of a parameter set for the whole period tends to produce poor prediction for the calm period. Nevertheless it is encouraging that the overall fit is good for most of the July data set, suggesting a certain robustness for the model over this summer period. Further work is in progress based on winter and spring data sets to assess the degree of variability of predictors over seasonal time-scales.
The MISO model is, of course, based on statistical analysis to produce an optimum prediction, and is not primarily designed to elucidate processes. However it is clear that the primary reason for its success is that it allows for "memory" in the suspended sediment signal whereas the LR method expects instantaneous response to two comparing hydro-

dynamic conditions. This "memory" is the result of suspended sediment taking longer to settle out of the water column than the time-scales of hydrodynamic variations, whether due to tides or storms. Prediction of these effect by fully process-based models will require much more sophistication than the simple LR or MISO models considered here.

Chapter 7

Multivariable Time Series Models for Sediment Dynamics

In this chapter, multivariable time series models of current velocity profile and suspended sediment concentration profile are presented. The current velocity and SPM dynamics are considered as unknown systems to be identified. The quantity relationships between them are investigated. The model structure, order determination problem are discussed and the model unknown parameter matrices are identified based upon the on-line recursive least squares identification method. The simulation results based on the real data collected from the Rufiji Delta, Tanzania show that these models are a good approach to data fitting and prediction.

7.1 Field Description and Data Collecting

The Rufiji delta in Tanzania contains the largest area of estuarine mangroves in East Africa: an area of 53000 hectares. The Rufiji river is located at latitude $7^o50'S$ within the tropics, (see Figure 7.1.1). In the months December until April the N-NE monsoon dominates causing abundant precipation, the S-SE monsoon is prevalent during May to October and causes significantly less precipitation. The deltaic plain formed at the Indian Ocean by the Rufiji river is approximately 23km wide and 70km long. Usual erosion and sedimentation patterns due to meander bend migration characteristic of deltaic estuaries

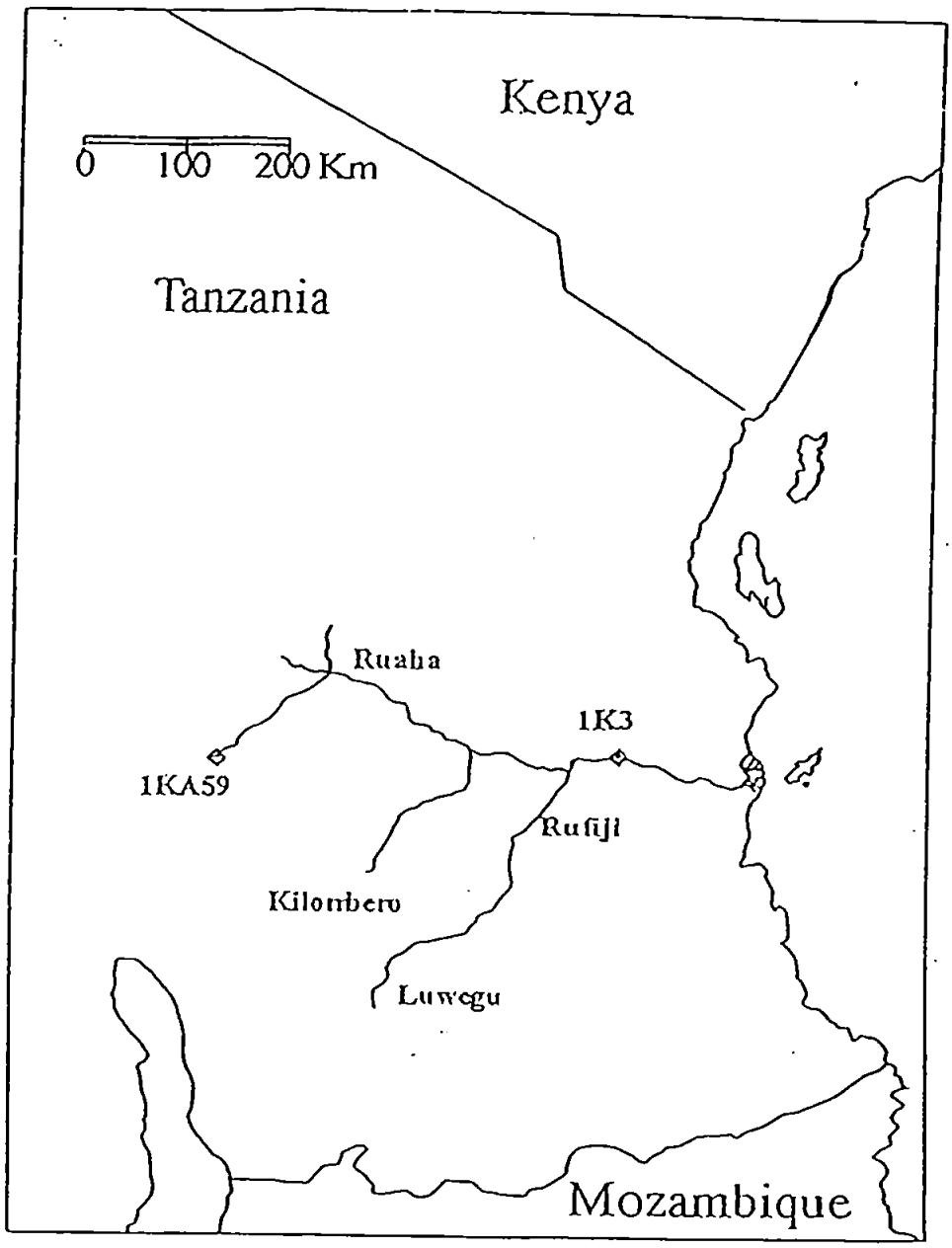


Figure 7.1.1. Location of Rufiji River on East African coast

can be witnessed throughout the delta. Tidal range at the estuary is approximately 4.3 meters and tidal observations show a symmetrical tidal curve at the river mouth. Sediments observed at the delta edge are sand, clays and silts varying in proportion and with high organic content and the data were collected by Fisher (1994).

7.1.1 Current Velocity Measurements

The current vane technique of Kjerfve and Medeiros (1989) was used and two current vanes were constructed using exactly the same materials and measurements described by Kjerfve and Medeiros, so that the same calibration equations could be used to correlate the vane deflection angles to current velocity. The vanes were lowered from the survey vessels on a graduated steel cable using a hand winch. Depth, deflection and current bearing were noted so that true depth (the vertical component of the measured depth) and longitudinal and lateral components of the current could be calculated. Deflection angles were kept within the recommended range for maximum accuracy by the use of interchangeable weights, so that the technique gives a 1% error in current velocity for a 1° error in deflection. As with tidal measurements observations were visually averaged to negate the effects of swell. The error in current measurement is ±3%. Errors in estuarine depth are 0.02m arising from inaccurate deflection observations and 0.1m arising from line reading errors (Fisher, 1994).

7.1.2 Suspended Sediment Sampling

The water sample was processed through a vacuum filter pump using 47mm diameter filter papers with guaranteed 0.45μm pore spaces. During filtration the samples were continually agitated, to ensure that the sediment remained in suspension. The volume of sample processed was dependent on concentration, but was typically in the range 50-500 ml. Sediment was removed by treating the filter paper with hydrogen peroxide and perchloric acid in the ratio of 1:1. The sediment was allowed to settle before the solution was decanted. The remaining sediment was then washed, dried and weighed. The use of this technique resulted in the destruction of organic matter within the sample, so that the resultant data is used for the calculation of transport of inorganic sediments only. It was found that sediment samples taken from the Rufiji river contained between 17.0 and 19.5% organics, with average of 17.8% (Fisher, 1994).

7.2 Current Velocity Model

In this section we focus on modelling the current velocity dynamics. The one reason we investigate the current velocity dynamics is that the most widely used sediment transport models are so-called cu-integral (concentration times velocity integral) type of models. The other reason is current velocity dynamics plays a very important part in sediment transport numerical model as we mentioned in Chapter 5 and Chapter 6. So current velocity is a very important variable to model when considering sediment transport. From Chapter 5, we know that the current velocity mainly is a function of its derivatives relevant to the direction (x, y, z) and some parameters, so in this section, the univariate model for current velocity profile is presented.

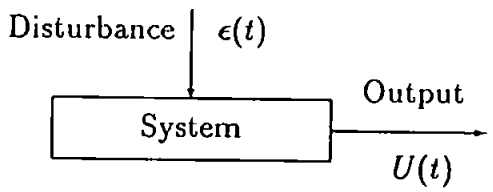


Fig 7.2.1 A Dynamic system with output $U(t)$ and disturbance $\epsilon(t)$ where t denotes time.

The distinguishing feature of a univariate time series current velocity model is that no attempt is made to relate $U(t)$ to other variables except the uncontrollable disturbance $\epsilon(t)$. The variations in $U(t)$ are 'explained' solely in terms of its own past or by its time and surrounding value, although of course if $U(t)$ is a vector dependent on space. The forecasts are then made by extrapolation. The statistical approach to forecasting is based on the construction of a model. The model defines a mechanism which is regarded as being capable of having produced the observations in question. Such a model of course when applied to the environment is of course invariably stochastic.

A single variable model for Tamar Estuary is presented in Chen and Dyke, 1995 and for the Holderness Coast in Chapter 6. Here, the current velocity profile (i.e. current velocity at different water depth) is taken as the output of the system and the model works in a similar way to the one dimensional vertical current velocity model, which is more realistic and a better description for the real system. We assume that the time series

model here that describes the current velocity is a discrete multivariable time-invariant stochastic linear system and can be represented by the following AR(p) or ARMA(p, q) model.

7.2.1 AR Model

1. Model Description

First, we assume that the current velocity profile model is a multivariable AR(p) model as follows:

$$A(z^{-1})U_n = w_n \tag{7.1}$$

where w_n is the system noise and the restriction on it are given in equations (3.66)-(3.67).

$$A(z^{-1}) = I_m + A_1 z^{-1} + ... + A_p z^{-p} \tag{7.2}$$

U_n and w_n are m-dimensional vectors instead of scalar as in the chapter 6. z^{-1} is a unit delay operator and $A_i, (i = 1, ..., p)$ is $m \times m$ unknown matrices to be estimated. I_m is an $m \times m$ unit matrix.
Set

$$\theta^T = [-A_1, ..., -A_p]_{m \times d} \tag{7.3}$$

$$x_n^T = [U_{n-1}^T, ..., U_{n-p}^T]_{1 \times d} \tag{7.4}$$

$$d = m \times p \tag{7.5}$$

here θ_n is the estimate of θ at time n and $[.]_{m \times d}$ and $[.]_{1 \times d}$ denote an $m \times d$ matrix and a $d-$dimensional row vector respectively.
It is easy to see that (7.1) also can be written as

$$U_n = \theta^T x_n + w_n \tag{7.6}$$

We construct the vertical profile variable vector of suspended sediment concentration or current velocity as the system output and it is a function of the water depth and time. Denote $U_n(h)$ as the current velocity at height h and time n (i.e. set water depth H, bottom $h = H$ and surface $h = 0$). So
$U_n^T = [U_n(h_1), U_n(h_2), U_n(h_3), ..., U_n(h_m)]_{1\times m}$. where $0 \leq h_1 \leq h_2 \leq ... \leq h_m \leq H$.

The recursive least squares algorithm of Chapter 3 (3.55–3.57) is used to identify the parameter matrix θ.

2. Order determination

Definition 7.1.
σ^2 is defined as variance of one-step prediction error of the model as follows:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} \|\tilde{U}_n\|^2 \tag{7.7}$$

where $\tilde{U}_n = U_n - \hat{U}_n$, U_n is the real data and $\hat{U}_n = x_n^T\theta_{n-1}$ is the one-step prediction of the model at time n respectively.

Definition 7.2.

$$MPVE = \max_{1\leq n\leq N} \|\tilde{U}_n\| \tag{7.8}$$

is defined the maximum one-step prediction vector error (MPVE) of the model.

Definition 7.3.

$$MPEE = \max_{\{1\leq n\leq N\}\bigcup\{1\leq j\leq m\}} |\tilde{U}_{n_j}| \tag{7.9}$$

is defined the maximum one-step prediction element error (MPEE) of the model, where $\tilde{U}_{n_j}$ is the $j-$th component of $\tilde{U}_n$.

Definition 7.4.

$$MPV = \max_{1\leq n\leq N} \|\tilde{\theta}_n\| \tag{7.10}$$

is defined the maximum parameter variation (MPV) of the model.

The approach in this chapter is to fit the model of progressively higher order, to calculate variance of one-step prediction error σ^2 for each value of order p, as well as to consider the MPVE, MPEE and MPV. The criterion is that if the addition of extra parameter matrices gives little improvement, we do not choose a higher order model.
The F-test results for our model candidates according to (4.11) and (4.12) are given as follows in Table 7.2.1:

TABLE 7.2.1. The order comparison of multivariable current velocity model

	MPVE	MPEE	MPV	σ^2
p=4	0.0327049	0.0455715	0.0718552	3.39857e-05
p=5	0.0326627	0.0456293	0.0508491	2.05273e-05
p=6	0.0326586	0.0456616	0.0597631	2.09884e-05

(i) Let AR(4)=$\mathcal{U}_1$, AR(5)=$\mathcal{U}_2$
$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{3.39857e-05-2.0527e-05}{2.0527e-05} = 239.306$ reject AR(4)

(ii) Let AR(5)=$\mathcal{U}_1$, AR(6)=$\mathcal{U}_2$
$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{2.0527e-05-2.09884e-05}{2.09884e-05} = -7.95074$ choose AR(5)

3. Simulation

Here, we choose a multivariable time series model which is a simpler form of (7.1). It follows the AR(5) model according to the order determination presented in Table 7.2.1:

$$U_n = AU_{n-1} + BU_{n-2} + CU_{n-3} + DU_{n-4} + EU_{n-5} + w_n \tag{7.11}$$

Set as:

$$\theta^T = [A, B, C, D, E] \tag{7.12}$$

$$x_n^T = [U_{n-1}, U_{n-2}, U_{n-3}, U_{n-4}, U_{n-5}] \tag{7.13}$$

In traditional models that seek to simulate the behaviour of dissolved and suspended matter in estuaries, the hydrodynamic equations are solved. Closure is imposed by a

model of turbulence. These models, once straightforward and simple can now be highly complex. The $k-\epsilon$ turbulence models can be modified to include two phases (sediment and fluid) and biology (the interaction of plants and animals with sediment). However, these equations become at once complicated and controversial. Since the mathematical representation of such a set of equations is written in finite difference form, ultimately it can be approximated to a difference equation (Bagchi and ten Brumelhuis, 1996) which would be similar to an equation such as (7.11). The fact that the so called 'constant matrices' in the equation do indeed remain constant is shown in the simulation that follow although they do show a very small variation with time (see Fig 7.2.12). These figures thus show the goodness of fit of the model to the data.

Returning to solving equation (7.11), the ELSM (3.55)–(3.57) is used (substitute U_n for y_n) in (3.55). From (7.5), here $p=5, m=10$ and $d=50$ and the time scale in here is $3\frac{3}{4}$ minutes per run. Since the initial value are needed, the first fifth data as our initial value we really start the model at n=6.

The computation procedure of ELSM is as follows:

(i) Construct x_n according to (7.13) ($n \geq 5$).
(ii) Select initial values of θ_5 and R_5.
(iii) Calculate K_n, R_n and θ_n according to (3.55)–(3.57) based on the K_{n-1}, R_{n-1} , θ_{n-1} and x_n ($n \geq 5$).

The simulation results are given as follows:
(i) The five parameter matrices are:

$$A = \begin{pmatrix} 1.499 & 0.227 & 0.351 & 0.106 & -0.105 & -0.047 & 0.074 & -0.090 & -0.083 & 0.019 \\ 0.289 & 1.260 & 0.331 & 0.065 & 0.091 & -0.126 & 0.074 & 0.061 & -0.084 & -0.012 \\ 0.234 & 0.187 & 1.185 & 0.153 & -0.069 & 0.006 & 0.156 & 0.012 & -0.044 & 0.106 \\ 0.065 & 0.050 & 0.318 & 0.908 & 0.152 & 0.168 & 0.142 & 0.006 & -0.008 & 0.148 \\ -0.116 & 0.117 & 0.108 & 0.131 & 1.250 & 0.143 & 0.124 & -0.008 & 0.101 & 0.078 \\ -0.048 & -0.114 & 0.191 & 0.092 & 0.125 & 1.312 & 0.250 & 0.100 & -0.061 & 0.099 \\ -0.100 & -0.045 & 0.223 & 0.004 & 0.045 & 0.216 & 1.297 & 0.303 & 0.055 & -0.114 \\ -0.171 & 0.108 & 0.152 & -0.022 & -0.018 & 0.168 & 0.446 & 1.230 & 0.108 & -0.070 \\ -0.071 & -0.029 & 0.152 & 0.016 & 0.100 & 0.000 & 0.121 & 0.107 & 1.291 & 0.349 \\ -0.055 & -0.100 & 0.155 & 0.112 & 0.017 & 0.078 & -0.151 & -0.089 & 0.258 & 1.668 \end{pmatrix}$$

$$B = \begin{pmatrix}
-0.339 & -0.441 & -0.298 & -0.137 & 0.068 & 0.001 & 0.042 & 0.083 & 0.063 & 0.009 \\
-0.446 & -0.066 & -0.267 & -0.155 & -0.091 & 0.057 & 0.025 & -0.104 & 0.040 & 0.067 \\
-0.338 & -0.290 & 0.171 & -0.196 & 0.024 & -0.056 & -0.058 & -0.065 & -0.011 & -0.082 \\
-0.143 & -0.147 & -0.128 & 0.216 & -0.118 & -0.169 & -0.105 & -0.101 & -0.077 & -0.158 \\
0.022 & -0.141 & 0.035 & -0.180 & 0.072 & -0.151 & -0.139 & -0.132 & -0.169 & -0.136 \\
0.019 & 0.042 & 0.009 & -0.229 & -0.119 & 0.000 & -0.271 & -0.247 & -0.026 & -0.100 \\
0.045 & 0.006 & -0.003 & -0.129 & -0.098 & -0.244 & -0.059 & -0.455 & -0.064 & 0.145 \\
0.100 & -0.048 & -0.028 & -0.092 & -0.105 & -0.213 & -0.421 & -0.128 & -0.106 & 0.127 \\
0.026 & 0.009 & 0.003 & -0.108 & -0.171 & -0.043 & -0.111 & -0.156 & -0.059 & -0.452 \\
0.034 & 0.055 & -0.041 & -0.130 & -0.074 & -0.072 & 0.148 & 0.166 & -0.389 & -0.559
\end{pmatrix}$$

$$C = \begin{pmatrix}
-0.080 & 0.055 & -0.003 & 0.025 & 0.100 & -0.043 & -0.054 & 0.001 & 0.044 & -0.044 \\
0.025 & -0.011 & -0.053 & 0.008 & -0.021 & 0.083 & 0.027 & -0.084 & 0.060 & -0.034 \\
-0.001 & -0.019 & 0.021 & -0.059 & 0.116 & 0.006 & -0.056 & -0.021 & 0.068 & -0.064 \\
0.018 & 0.006 & -0.073 & 0.075 & -0.015 & -0.054 & -0.027 & 0.018 & 0.082 & -0.048 \\
0.066 & -0.087 & 0.065 & -0.046 & -0.107 & -0.004 & -0.002 & 0.063 & 0.034 & 0.010 \\
-0.019 & 0.064 & -0.003 & -0.074 & 0.013 & -0.118 & -0.029 & 0.039 & 0.150 & -0.061 \\
-0.007 & 0.031 & -0.038 & 0.012 & 0.065 & 0.016 & -0.027 & -0.059 & 0.025 & -0.038 \\
0.048 & -0.063 & -0.033 & 0.051 & 0.099 & 0.056 & -0.081 & -0.041 & -0.005 & -0.051 \\
-0.012 & -0.008 & 0.010 & 0.057 & 0.029 & 0.117 & -0.054 & -0.060 & -0.061 & -0.023 \\
0.023 & -0.006 & -0.033 & 0.003 & 0.082 & -0.024 & -0.030 & -0.023 & 0.061 & -0.077
\end{pmatrix}$$

$$D = \begin{pmatrix}
-0.060 & 0.109 & -0.001 & 0.059 & 0.015 & -0.019 & -0.021 & -0.014 & 0.023 & -0.021 \\
0.078 & -0.069 & -0.021 & 0.063 & 0.019 & 0.003 & -0.009 & -0.011 & 0.036 & -0.023 \\
0.052 & 0.043 & -0.132 & 0.033 & 0.029 & 0.007 & -0.022 & 0.005 & 0.037 & -0.004 \\
0.031 & 0.030 & -0.055 & -0.049 & 0.006 & 0.008 & 0.000 & 0.025 & 0.039 & 0.012 \\
0.016 & 0.003 & -0.032 & 0.031 & -0.112 & 0.010 & 0.021 & 0.054 & 0.040 & 0.021 \\
-0.002 & 0.011 & -0.052 & 0.055 & 0.001 & -0.119 & 0.033 & 0.059 & 0.045 & 0.010 \\
0.002 & -0.002 & -0.062 & 0.045 & 0.033 & 0.030 & -0.087 & 0.056 & 0.018 & -0.012 \\
0.010 & -0.039 & -0.028 & 0.041 & 0.066 & 0.039 & 0.019 & -0.060 & 0.014 & -0.018 \\
-0.005 & -0.004 & -0.035 & 0.042 & 0.050 & 0.034 & 0.003 & 0.017 & -0.092 & 0.064 \\
0.000 & 0.004 & -0.040 & 0.030 & 0.032 & 0.017 & -0.011 & -0.043 & 0.076 & -0.026
\end{pmatrix}$$

$$E = \begin{pmatrix} -0.034 & 0.061 & -0.098 & 0.068 & -0.136 & 0.059 & 0.028 & 0.011 & -0.101 & 0.073 \\ 0.066 & -0.140 & -0.013 & 0.105 & -0.045 & -0.066 & -0.032 & 0.116 & -0.095 & 0.034 \\ -0.066 & 0.081 & -0.276 & 0.104 & -0.115 & 0.003 & 0.051 & 0.051 & -0.087 & 0.073 \\ 0.036 & 0.051 & -0.050 & -0.166 & -0.017 & 0.013 & 0.047 & 0.049 & -0.083 & 0.082 \\ -0.005 & 0.113 & -0.147 & 0.072 & -0.137 & -0.010 & 0.056 & 0.004 & -0.056 & 0.068 \\ 0.035 & 0.003 & -0.128 & 0.158 & -0.008 & -0.131 & 0.076 & 0.037 & -0.148 & 0.084 \\ 0.040 & 0.017 & -0.075 & 0.035 & -0.037 & -0.033 & -0.098 & 0.154 & -0.080 & 0.055 \\ 0.000 & 0.041 & -0.008 & -0.033 & -0.034 & -0.068 & 0.110 & -0.037 & -0.052 & 0.052 \\ 0.056 & 0.040 & -0.101 & -0.056 & -0.003 & -0.117 & 0.092 & 0.096 & -0.157 & 0.120 \\ -0.023 & 0.088 & -0.063 & 0.003 & -0.081 & -0.001 & 0.078 & -0.024 & -0.031 & 0.013 \end{pmatrix}$$

ii)

TABLE 7.2.2 The mode of the parameter matrices in AR(5) current velocity model

	A	B	C	D	E
p=5	2.029754	0.955137	0.335718	0.239755	0.624578

(iii)

MPEE=0.0326627 m/s

MPVE=0.0456293 m/s

MPV=0.0508491

σ^2=2.05273 e-05

(iv) Figures 7.2.2–7.2.11 show the simulation of the current velocity dynamics at different depth and Figure 7.2.12. gives the norm of the parameter matrix error dynamics in the current velocity model.

(v) The mode and the element of A, B are comparatively large and the ones of C, D, E are comparatively small which shows that the more recent the time, the more effect is there on the current variation.

(vi) Since A is strongly diagonally dominant, we can say that the larger the distance between given layers, the less effect is there on the layer variation.

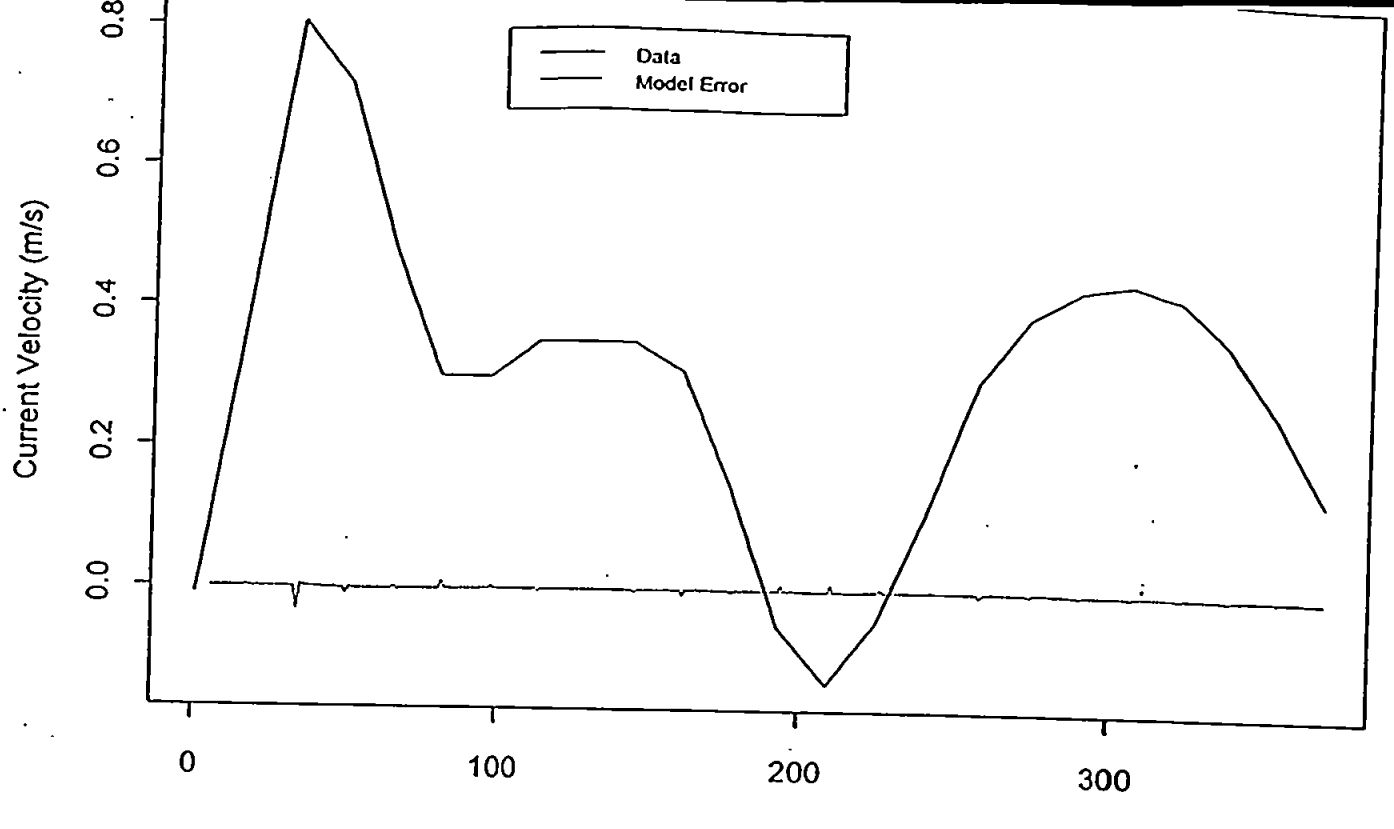


Figure 7.2.2. Plot of AR(5) Current Velocity Model Error vs Data at h=0.05H

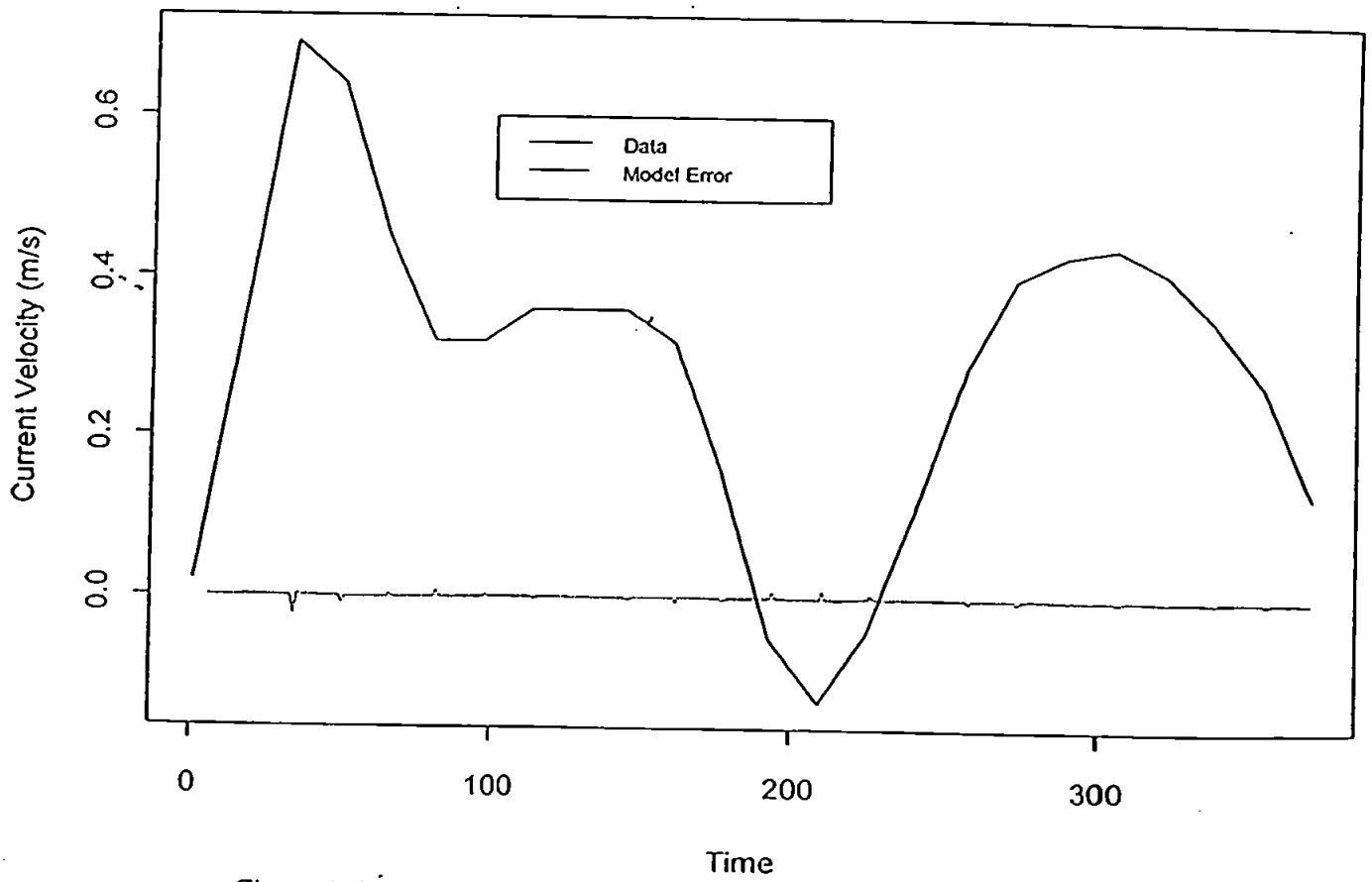


Figure 7.2.3. Plot of AR(5) Current Velocity Model Error vs Data at h=0.10H

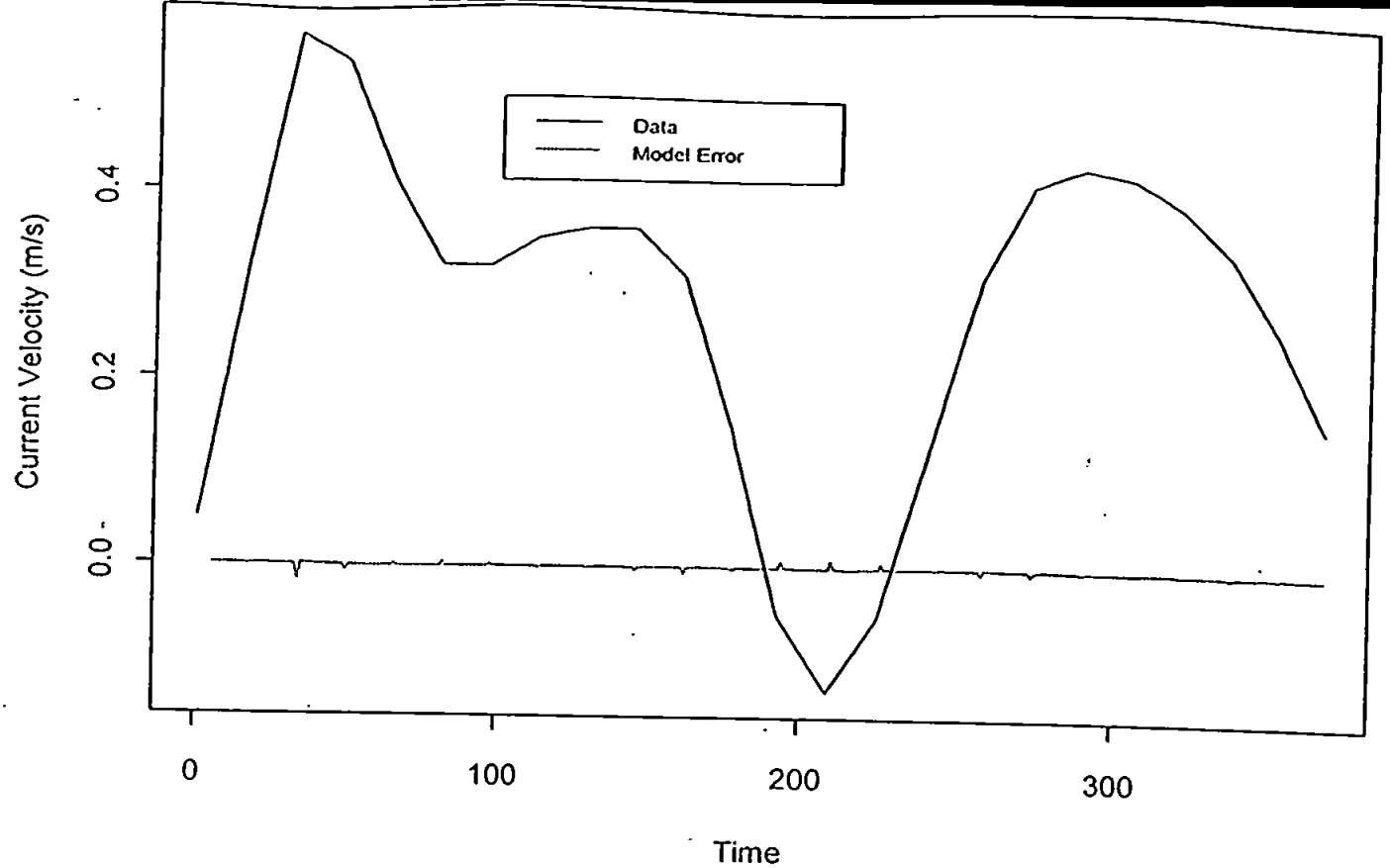


Figure 7.2.4. Plot of AR(5) Current Velocity Model Error vs Data at h=0.20H

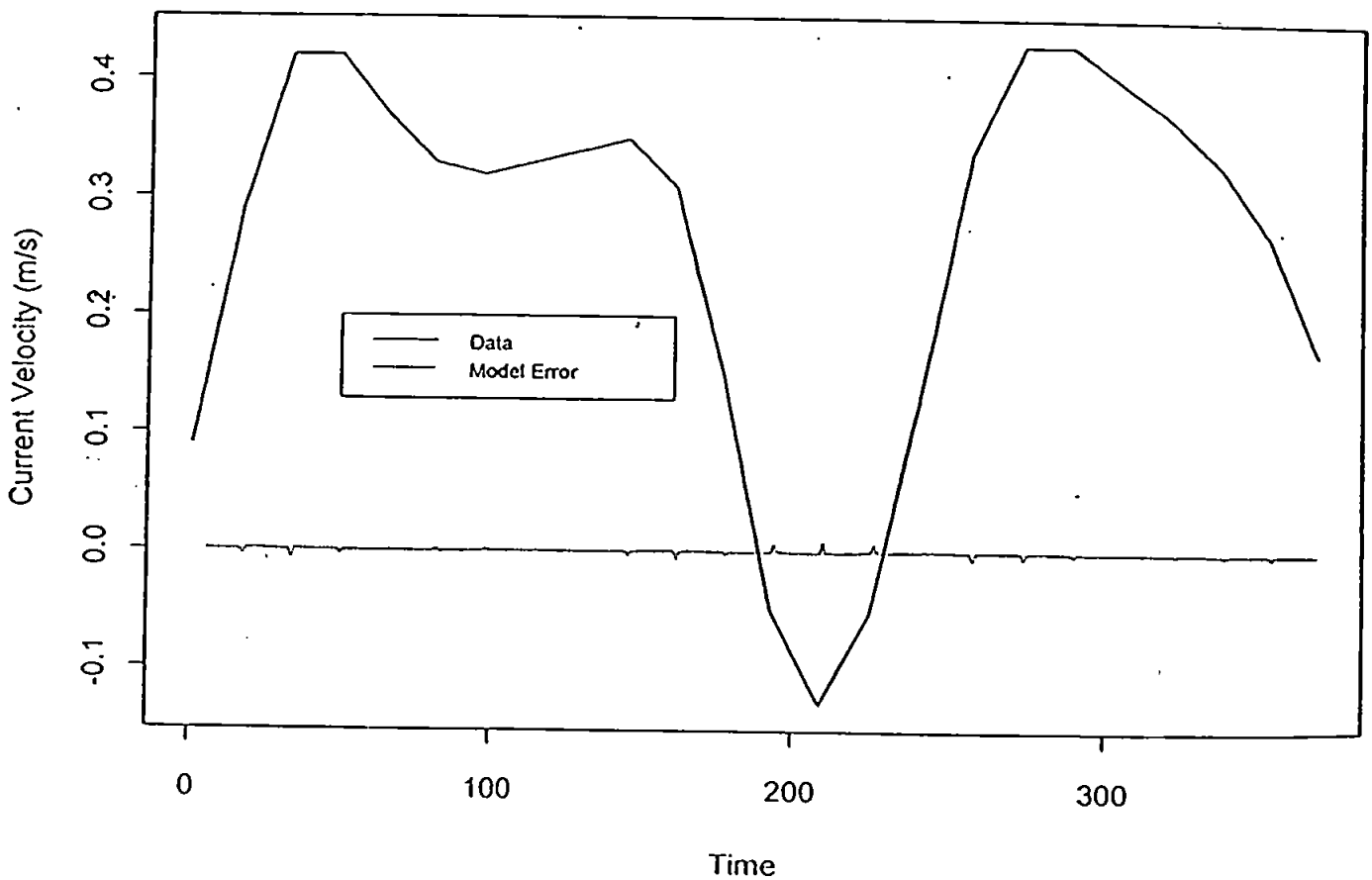


Figure 7.2.5. Plot of AR(5) Current Velocity Model Error vs Data at h=0.30H

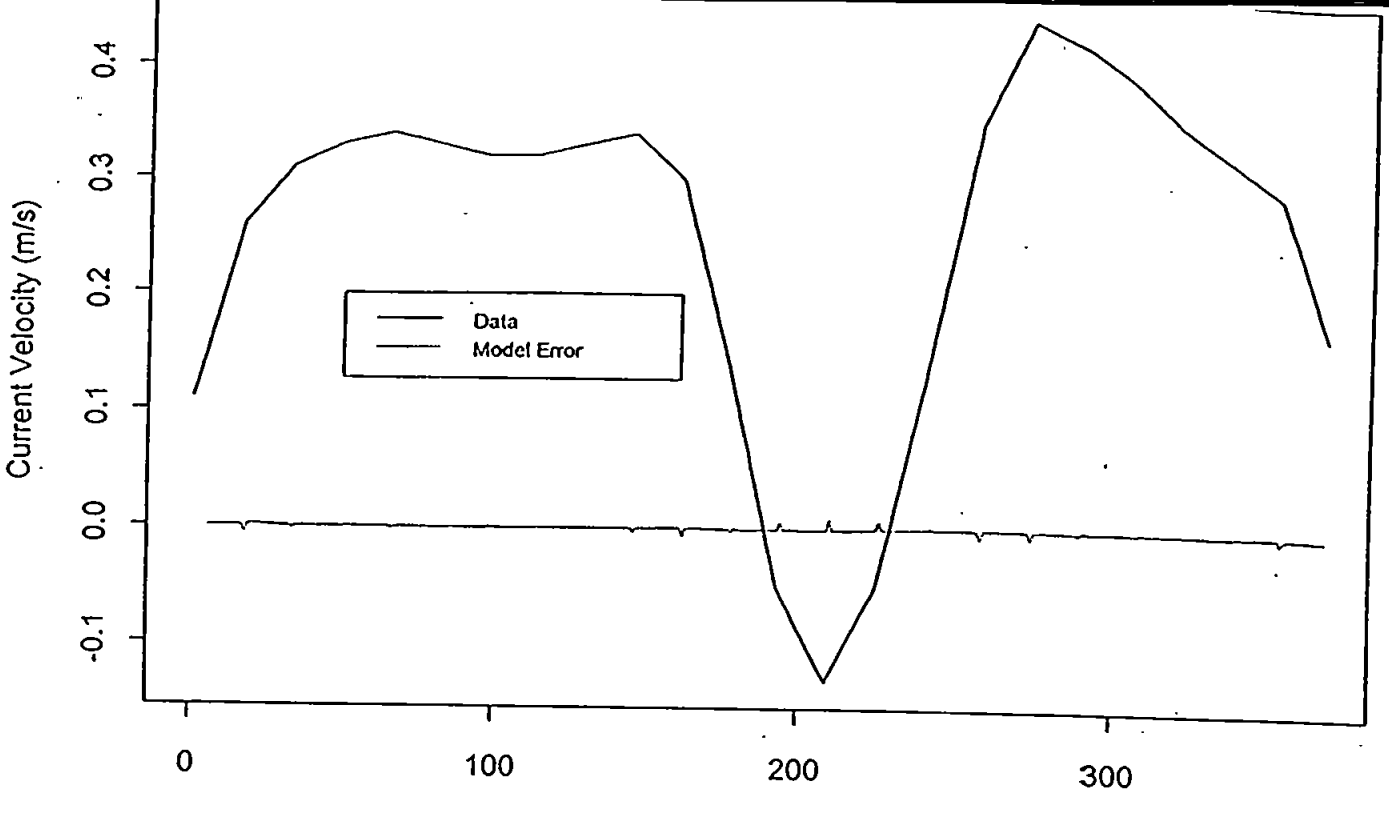


Figure 7.2.6. Plot of AR(5) Current Velocity Model Error vs Data at h=0.40H

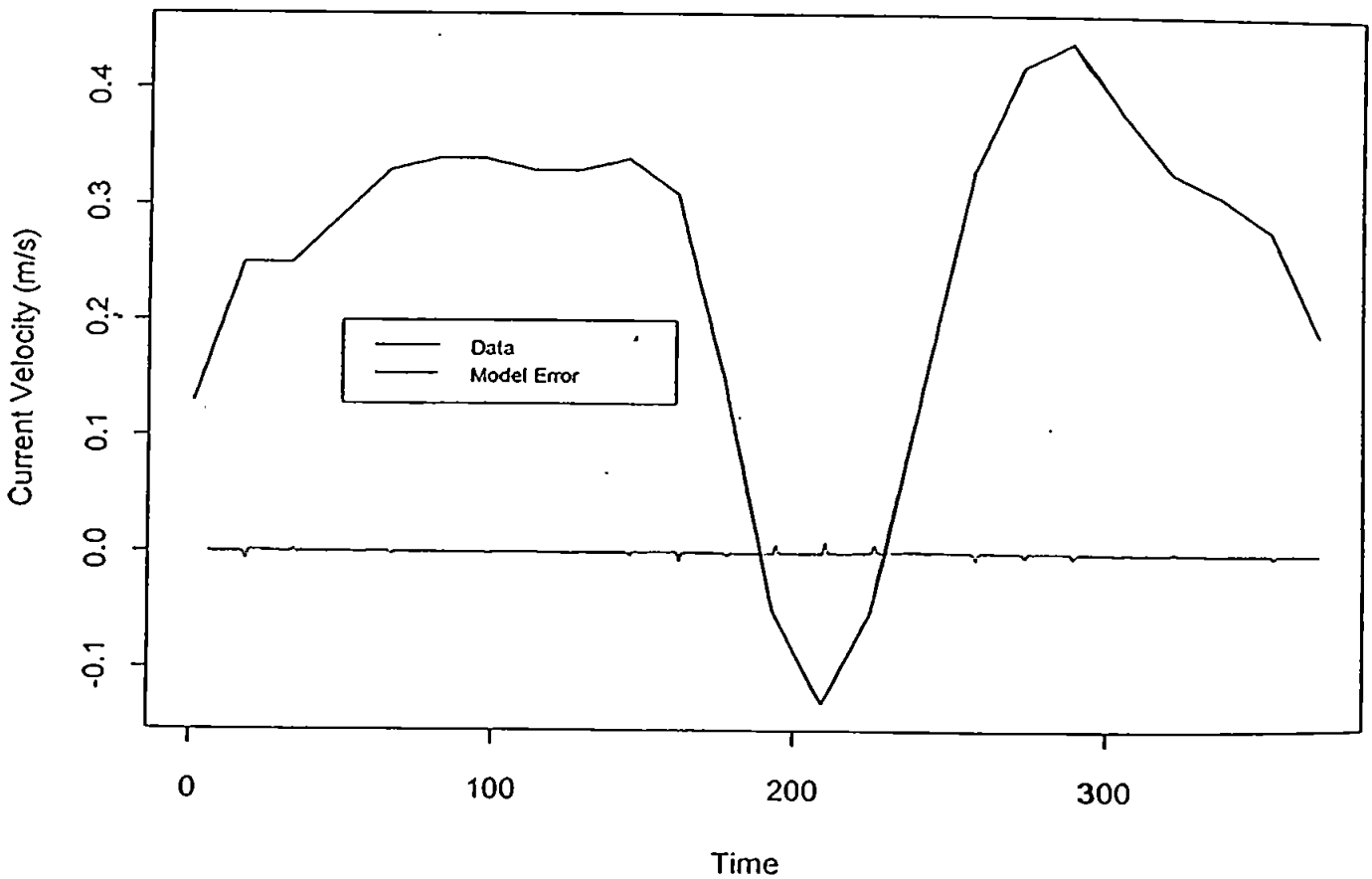


Figure 7.2.7. Plot of AR(5) Current Velocity Model Error vs Data at h=0.50H

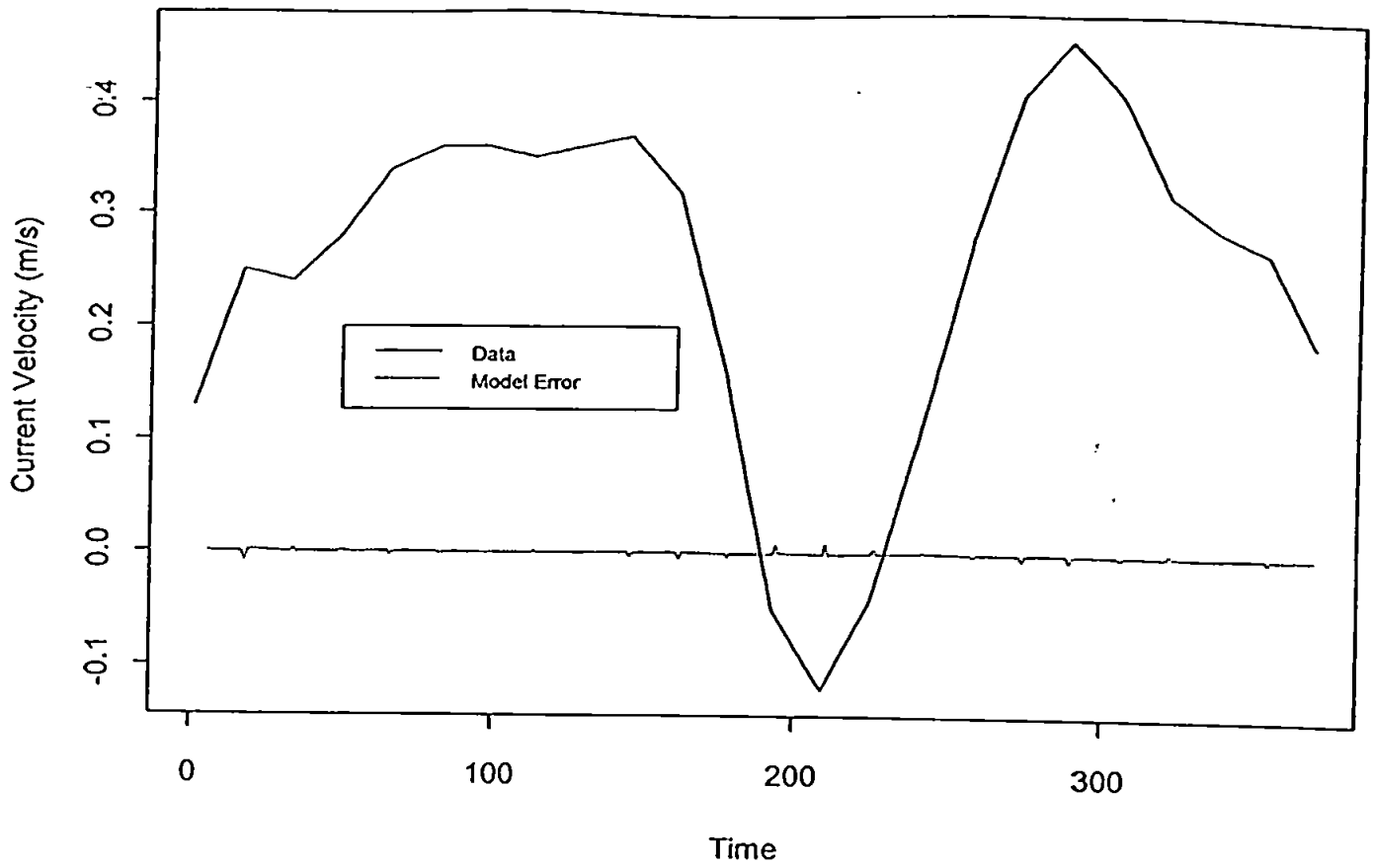


Figure 7.2.8. Plot of AR(5) Current Velocity Model Error vs Data at h=0.60H

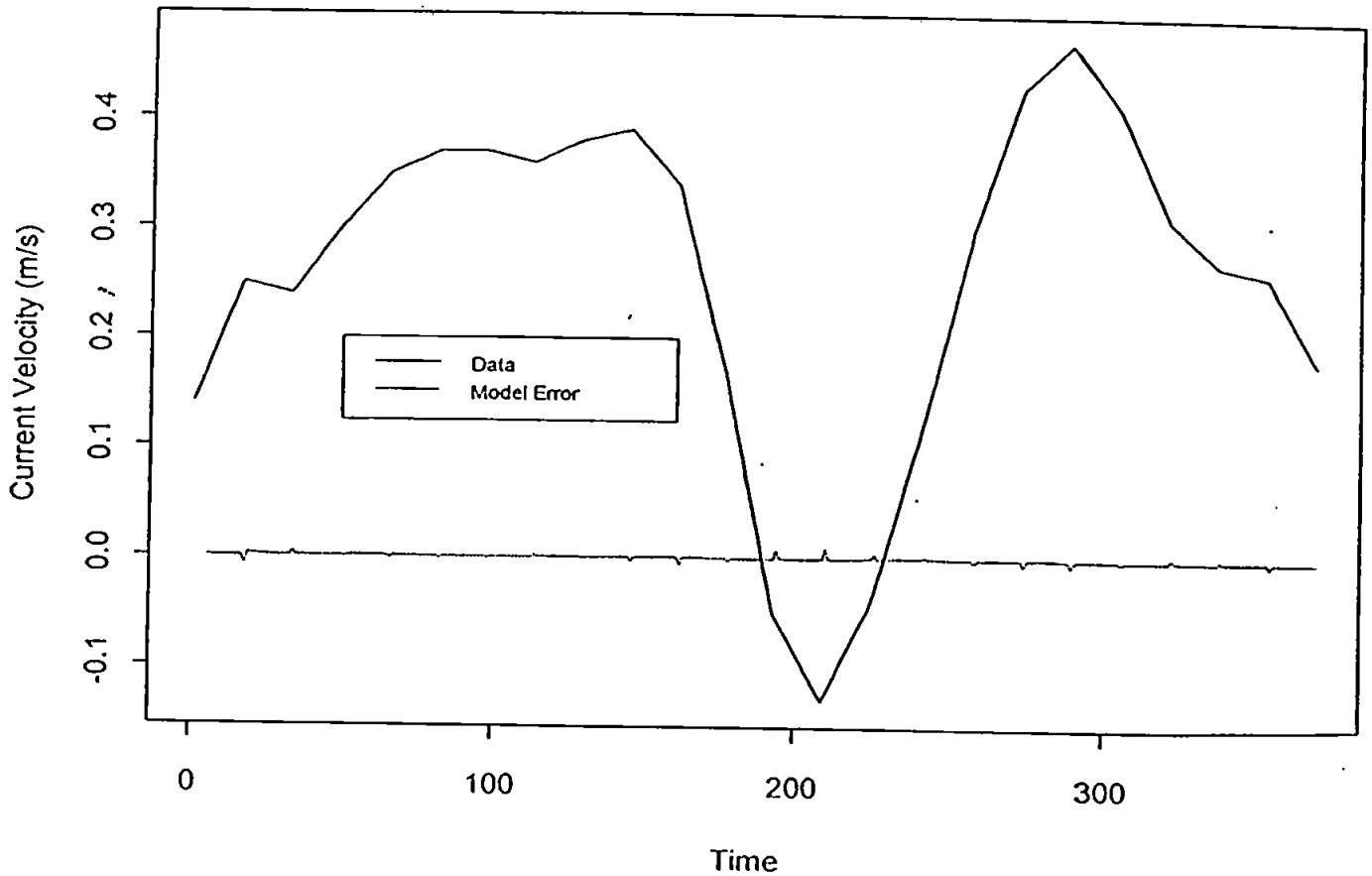


Figure 7.2.9. Plot of AR(5) Current Velocity Model Error vs Data at h=0.70H

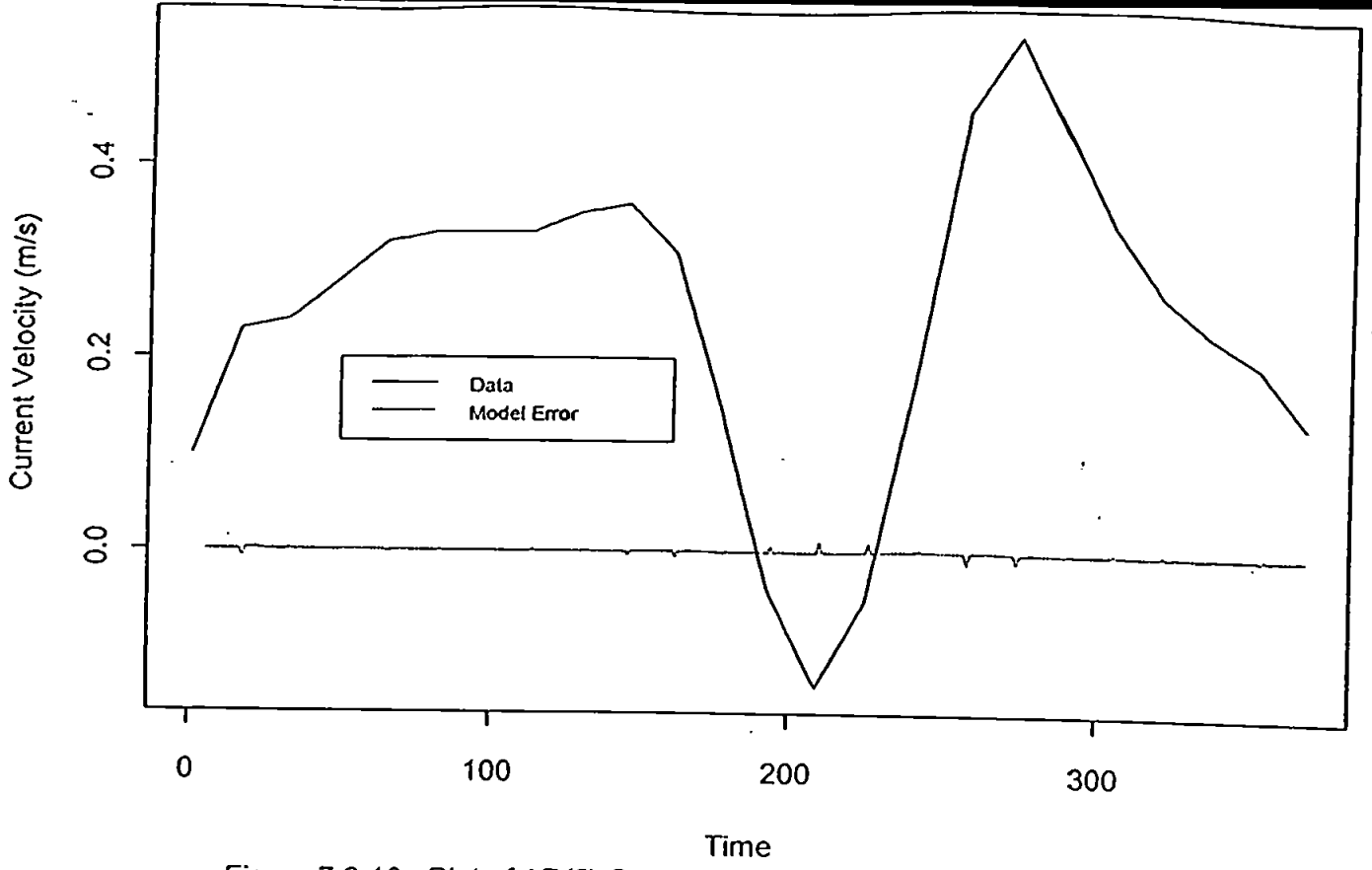


Figure 7.2.10. Plot of AR(5) Current Velocity Model Error vs Data at h=0.80H

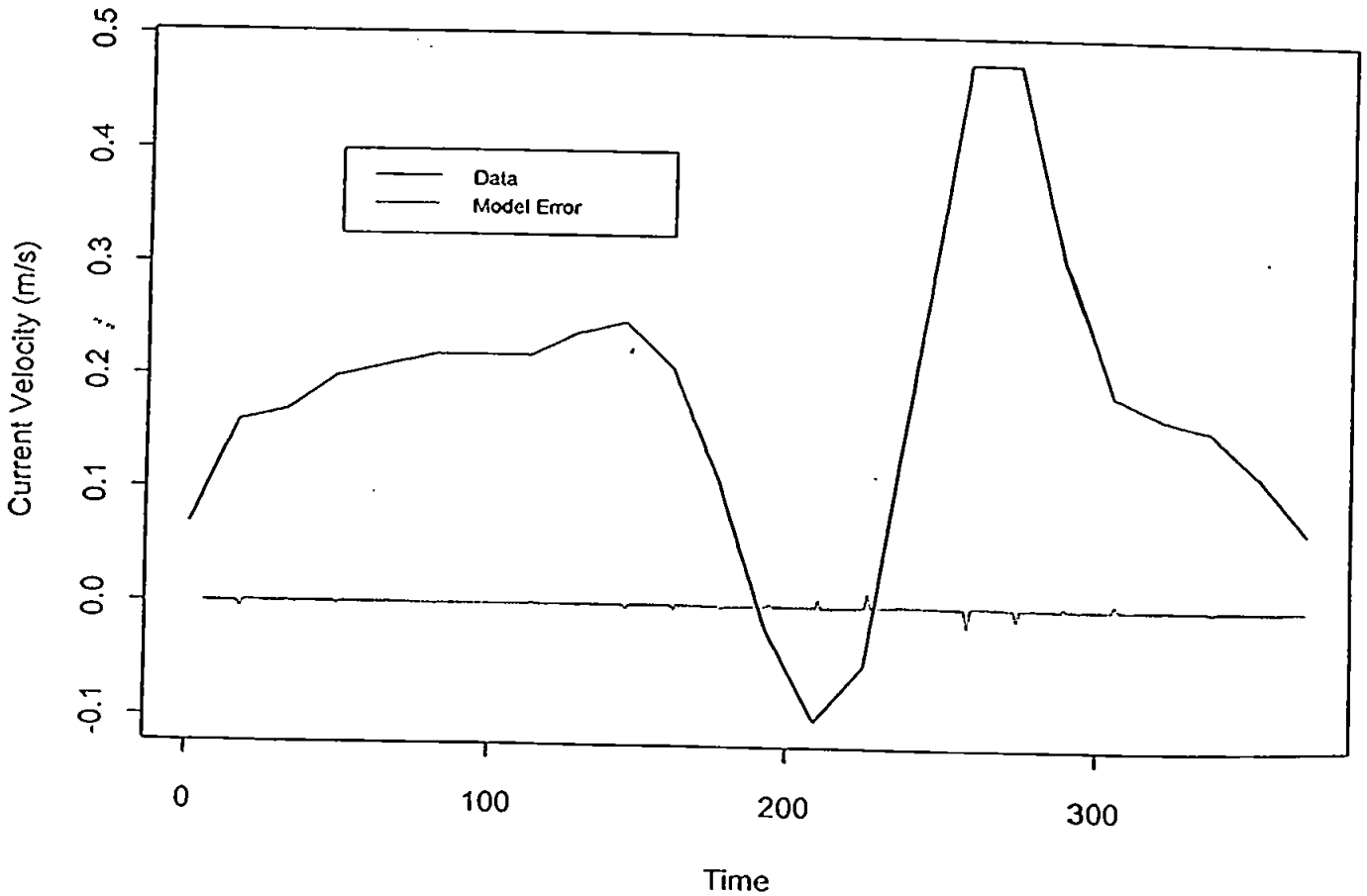


Figure 7.2.11. Plot of AR(5) Current Velocity Model Error vs Data at h=0.90H

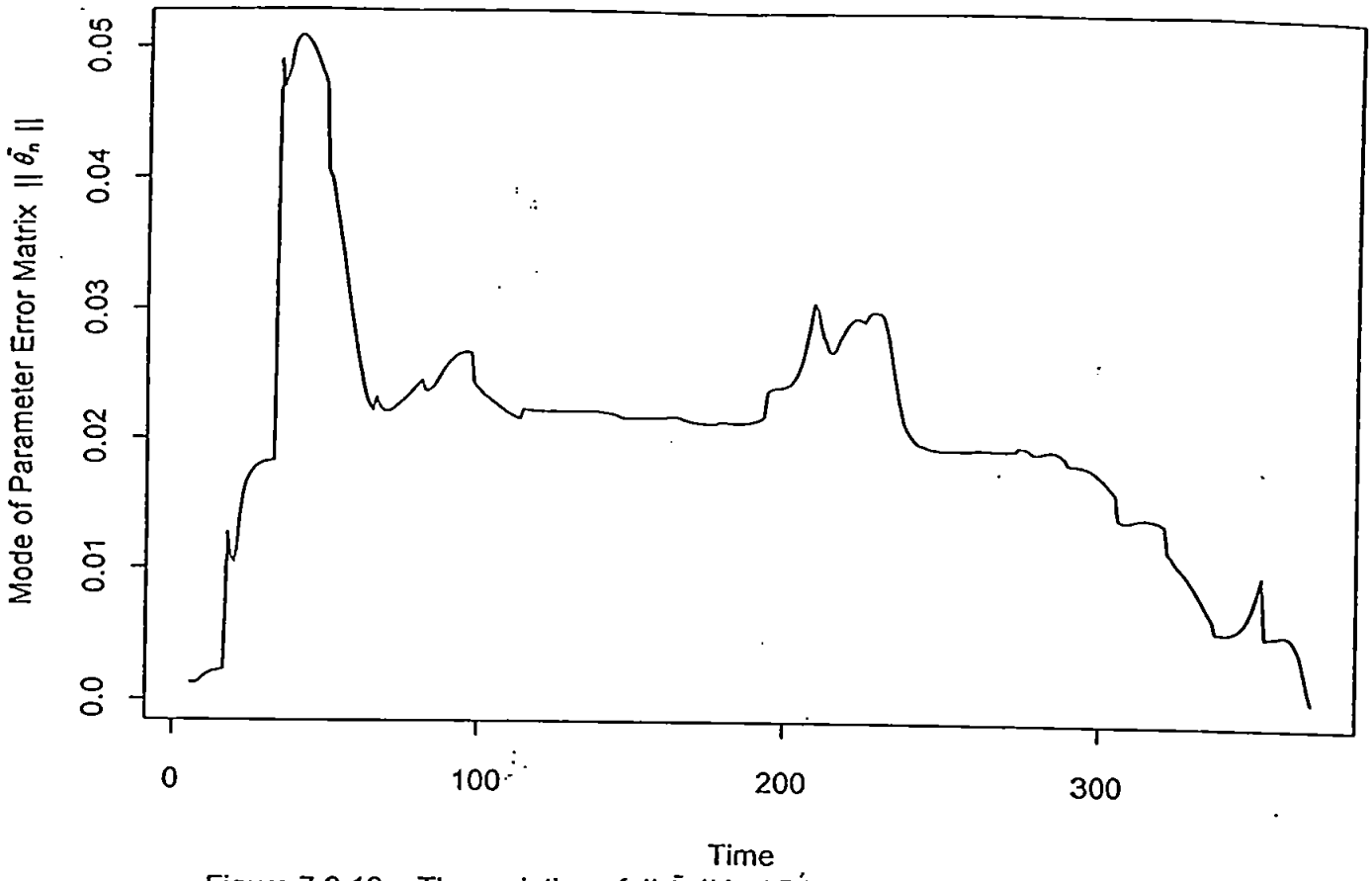


Figure 7.2.12. The variation of || $\tilde{\theta}_n$ || in AR(5) Current Velocity Model with time

7.2.2 ARMA Model

In this section, the more general system with the correlated system noise is considered, i.e., the multivariable ARMA model for current velocity profile are constructed.

1. Model Description

Now, we assume that the current velocity profile model is a multivariable ARMA(p,q) model as follows:

$$A(z^{-1})U_n = C(z^{-1})w_n \tag{7.14}$$

where w_n is the system noise and the restriction on it is given in (3.66) and (3.67).

$$A(z^{-1}) = I_m + A_1 z^{-1} + ... + A_p z^{-p}$$
$$C(z^{-1}) = I_m + C_1 z^{-1} + ... + C_q z^{-q}$$

U_n and w_n are m-dimensional vectors, z^{-1} is a unit delay operator and $A_i, C_j (i = 1,...,p; j = 1,...,q)$ is $m \times m$ unknown matrices to be estimated. I_m is an $m \times m$ unit matrix.
Set

$$\theta^T = [-A_1,...,-A_p, C_1,...,C_q]_{m\times d} \tag{7.15}$$

$$x_n^T = [U_{n-1}^T,...,U_{n-p}^T, e_{n-1}^T,...,e_{n-q}^T]_{1\times d} \tag{7.16}$$

$$d = m \times p \tag{7.17}$$

$$e_n = U_n - \theta_n^T x_n \tag{7.18}$$

here θ_n is the estimate of θ at time n and $[.]_{m\times d}$ and $[.]_{1\times d}$ denote an $m \times d$ matrix and a $d-$dimensional row vector respectively.
It is easy to see that (7.14) also can be written as

$$U_n = \theta^T x_n + C(z^{-1})w_n + e_n - C(z^{-1})e_n \tag{7.19}$$

We construct the vertical profile variable vector of suspended sediment concentration or current velocity as the system output and it is a function of the water depth and time.

The $U_n(h)$ is defined as the last subsection and
$U_n^T = [U_n(h_1), U_n(h_2), U_n(h_3), ..., U_n(h_m)]_{1\times m}$. where $0 \le h_1 \le h_2 \le ... \le h_m \le H$.

In order to identify the system parameter matrix θ, we make use of the recursive algorithms (3.55)-(3.57).

2. Order determination

The order determination method here is similar to that in the last subsection and the F-test results for our model candidates according to (4.11) and (4.12) are given as following in Table 7.2.3:

TABLE 7.2.3. The order comparison of current velocity model

	MPVE	MPEE	MPV	σ^2
p=2, q=1	0.0288694	0.0403688	0.00636337	2.91657e-05
p=3, q=1	0.028119	0.0393385	0.0061354	1.77705e-05
p=4, q=1	0.027965	0.0390262	0.00605504	1.83553e-05

(i) Let ARMA(2,1)=$\mathcal{U}_1$, ARMA(3,1)=$\mathcal{U}_2$
$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{2.91657e-05-1.77705e-05}{1.77705e-05} = 234.0535$ reject ARMA(2,1)

(ii) Let ARMA(3,1)=$\mathcal{U}_1$, ARMA(4,1)=$\mathcal{U}_2$
$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{1.77705e-05-1.83553e-05}{1.83553e-05} = -11.6289$ choose ARMA(3,1)

3. Simulation

According to the order determination presented in Table 7.2.3., we choose a dimensional multivariable time series model which is a simpler form of (7.14) as following ARMA(3,1) model:

$$U_n = A_1 U_{n-1} + A_2 U_{n-2} + A_3 U_{n-3} + w_n + C_1 w_{n-1} \tag{7.20}$$

Set as:

$$\theta^T = [A_1, A_2, A_3, C_1] \tag{7.21}$$

$$x_n^T = [U_{n-1}^T, U_{n-2}^T, U_{n-3}^T, e_{n-1}^T]. \tag{7.22}$$

In order to solve equation (7.20), the ELSM (3.55)–(3.57) is used (substitute U_n for y_n) in (3.55). From (7.20), here $p = 3, q = 1, m = 10$ and $d = 40$ and the time scale in here is $3\frac{3}{4}$ minutes per run. Since the initial value are need, the first third data as our initial value we really start the model at n=4.

The computation procedure of ELSM is as follows:

(i) Construct x_n according to (7.22) and (7.18) ($n \geq 3$).
(ii) Select initial values of θ_3 and R_3.
(iii) Calculate K_n, R_n and θ_n according to (3.55)–(3.57) based on the K_{n-1}, R_{n-1} , θ_{n-1} and x_n ($n \geq 3$).

The simulation results are given as follows:
(i) The five parameter matrices are:

$$A_1 = \begin{pmatrix} 0.989 & 0.249 & 0.147 & 0.061 & -0.055 & -0.076 & 0.039 & -0.014 & -0.035 & 0.033 \\ 0.240 & 0.791 & 0.172 & 0.142 & 0.043 & -0.029 & 0.039 & -0.009 & 0.001 & -0.053 \\ 0.193 & 0.175 & 0.640 & 0.147 & 0.022 & 0.045 & 0.076 & -0.043 & 0.032 & 0.028 \\ 0.015 & 0.062 & 0.201 & 0.616 & 0.185 & 0.142 & 0.076 & -0.058 & 0.041 & 0.054 \\ -0.109 & 0.056 & 0.081 & 0.254 & 0.805 & 0.132 & 0.112 & -0.060 & 0.028 & 0.044 \\ -0.107 & -0.051 & 0.109 & 0.230 & 0.139 & 0.726 & 0.197 & 0.064 & -0.002 & 0.026 \\ -0.040 & -0.039 & 0.095 & 0.062 & 0.093 & 0.201 & 0.747 & 0.267 & 0.041 & -0.119 \\ -0.049 & 0.001 & 0.085 & -0.042 & 0.043 & 0.148 & 0.345 & 0.768 & 0.064 & -0.056 \\ 0.010 & -0.026 & 0.033 & -0.027 & 0.045 & -0.002 & 0.121 & 0.157 & 0.751 & 0.356 \\ 0.035 & -0.022 & -0.056 & -0.001 & -0.046 & -0.033 & -0.055 & -0.002 & 0.302 & 1.131 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0.352 & 0.002 & -0.012 & 0.038 & -0.015 & -0.021 & 0.023 & -0.006 & -0.018 & 0.026 \\ 0.008 & 0.352 & -0.020 & 0.044 & -0.022 & -0.020 & 0.035 & -0.024 & -0.017 & 0.016 \\ 0.008 & -0.006 & 0.323 & -0.001 & 0.001 & 0.001 & 0.043 & -0.017 & 0.007 & 0.007 \\ 0.005 & -0.019 & 0.001 & 0.345 & 0.004 & -0.018 & 0.034 & -0.006 & -0.005 & 0.004 \\ -0.019 & 0.007 & -0.004 & 0.018 & 0.313 & -0.012 & 0.046 & -0.015 & -0.018 & 0.013 \\ -0.017 & 0.014 & -0.006 & 0.038 & -0.015 & 0.313 & 0.026 & -0.012 & -0.006 & -0.004 \\ -0.024 & -0.017 & 0.005 & -0.022 & -0.002 & 0.002 & 0.341 & 0.002 & -0.021 & 0.012 \\ -0.009 & -0.005 & 0.038 & -0.054 & 0.021 & 0.006 & 0.030 & 0.328 & -0.018 & 0.024 \\ 0.016 & -0.011 & 0.013 & -0.040 & 0.020 & -0.002 & 0.012 & 0.006 & 0.301 & 0.028 \\ -0.005 & 0.029 & -0.007 & 0.044 & -0.017 & -0.002 & 0.030 & -0.024 & 0.013 & 0.336 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} -0.308 & -0.285 & -0.193 & 0.001 & 0.021 & 0.048 & 0.034 & 0.018 & -0.023 & -0.006 \\ -0.261 & -0.106 & -0.229 & -0.069 & -0.088 & -0.010 & 0.050 & -0.012 & -0.034 & 0.080 \\ -0.164 & -0.192 & 0.008 & -0.155 & -0.043 & -0.059 & -0.002 & -0.009 & -0.038 & -0.024 \\ 0.005 & -0.094 & -0.199 & 0.068 & -0.191 & -0.190 & -0.0174 & 0.034 & -0.057 & -0.038 \\ 0.091 & -0.023 & -0.083 & -0.215 & -0.167 & -0.172 & -0.033 & 0.015 & -0.062 & -0.009 \\ 0.077 & 0.075 & -0.124 & -0.153 & -0.178 & -0.095 & -0.151 & -0.098 & -0.006 & -0.009 \\ 0.061 & 0.045 & -0.049 & -0.087 & -0.093 & -0.196 & -0.060 & -0.264 & -0.078 & 0.152 \\ 0.048 & -0.001 & -0.011 & -0.075 & -0.016 & -0.154 & -0.298 & -0.123 & -0.115 & 0.093 \\ 0.018 & 0.000 & -0.013 & -0.060 & -0.008 & -0.004 & -0.083 & -0.133 & -0.145 & -0.317 \\ -0.079 & 0.057 & 0.017 & 0.067 & -0.015 & 0.012 & 0.112 & -0.052 & -0.297 & -0.475 \end{pmatrix}$$

$$C_1 = \begin{pmatrix} 0.604 & 0.064 & 0.091 & -0.031 & -0.063 & -0.052 & 0.037 & -0.026 & -0.012 & 0.025 \\ 0.059 & 0.521 & 0.112 & 0.029 & 0.013 & -0.052 & -0.027 & -0.001 & 0.003 & -0.039 \\ 0.044 & 0.076 & 0.449 & 0.114 & -0.037 & 0.036 & 0.035 & -0.054 & -0.011 & -0.009 \\ -0.016 & 0.017 & 0.132 & 0.376 & 0.109 & 0.110 & 0.013 & -0.103 & 0.003 & 0.000 \\ -0.064 & 0.022 & 0.023 & 0.105 & 0.577 & 0.011 & 0.022 & -0.091 & 0.025 & -0.010 \\ -0.025 & -0.068 & 0.077 & 0.081 & 0.047 & 0.486 & 0.087 & -0.029 & -0.020 & 0.010 \\ -0.000 & -0.048 & 0.021 & 0.004 & -0.009 & 0.085 & 0.455 & 0.128 & 0.029 & -0.067 \\ -0.002 & -0.010 & 0.020 & -0.040 & -0.052 & 0.031 & 0.177 & 0.496 & 0.045 & -0.022 \\ -0.013 & 0.018 & -0.006 & 0.001 & -0.009 & -0.053 & 0.067 & 0.053 & 0.489 & 0.120 \\ 0.049 & -0.016 & -0.039 & -0.026 & 0.004 & -0.003 & -0.039 & -0.001 & 0.064 & 0.664 \end{pmatrix}$$

ii)

TABLE 7.2.4. The mode of the parameter matrices in ARMA(3,1) current velocity model

	A_1	A_2	A_3	C_1
ARMA(3,1)	1.35035	0.417748	0.699889	0.752131

(iii)

MPEE=0.028119 m/s

MPVE=0.0393385 m/s

MPV= 0.0061354

σ^2=1.77705 e-05

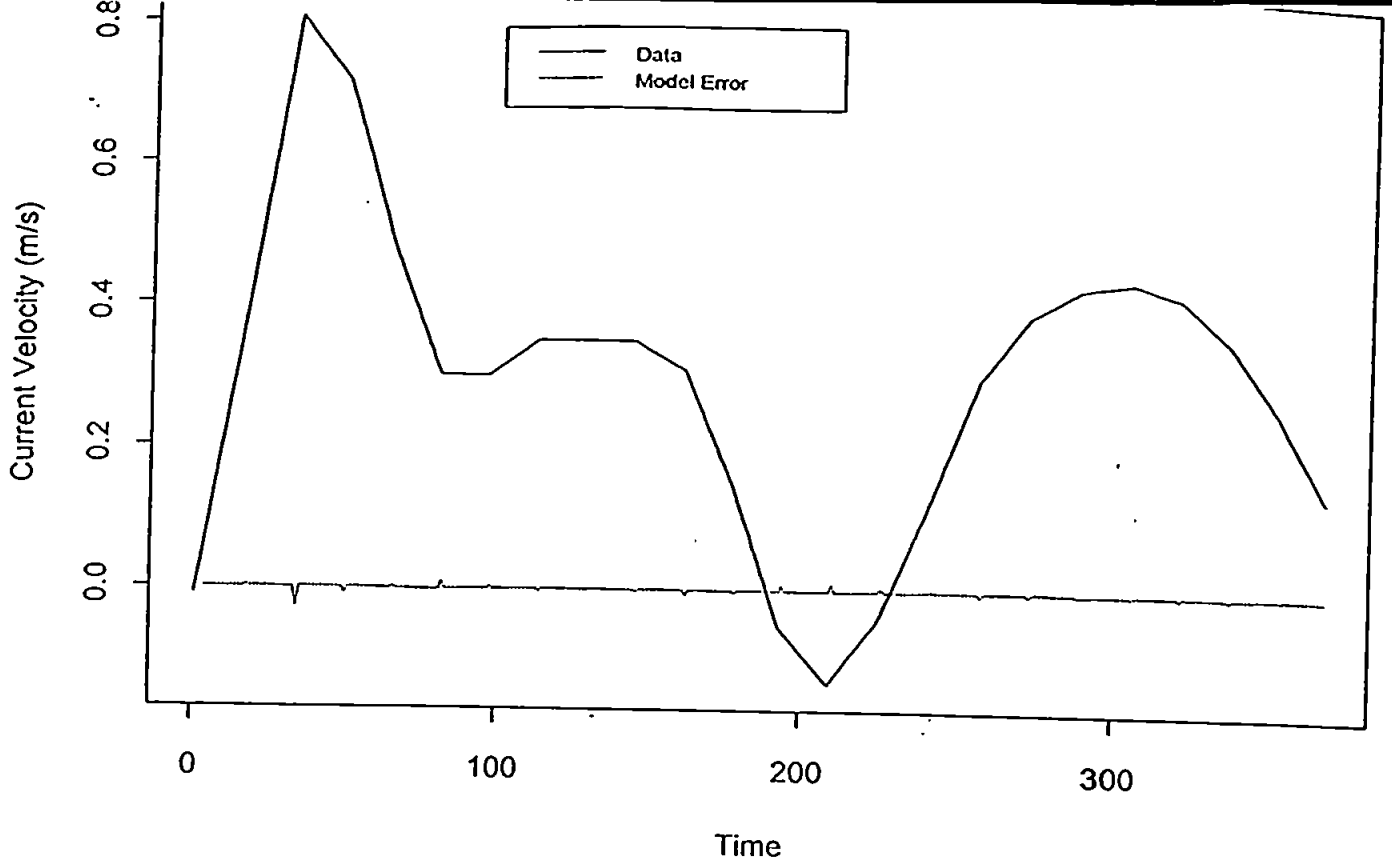


Figure 7.2.13. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.05H

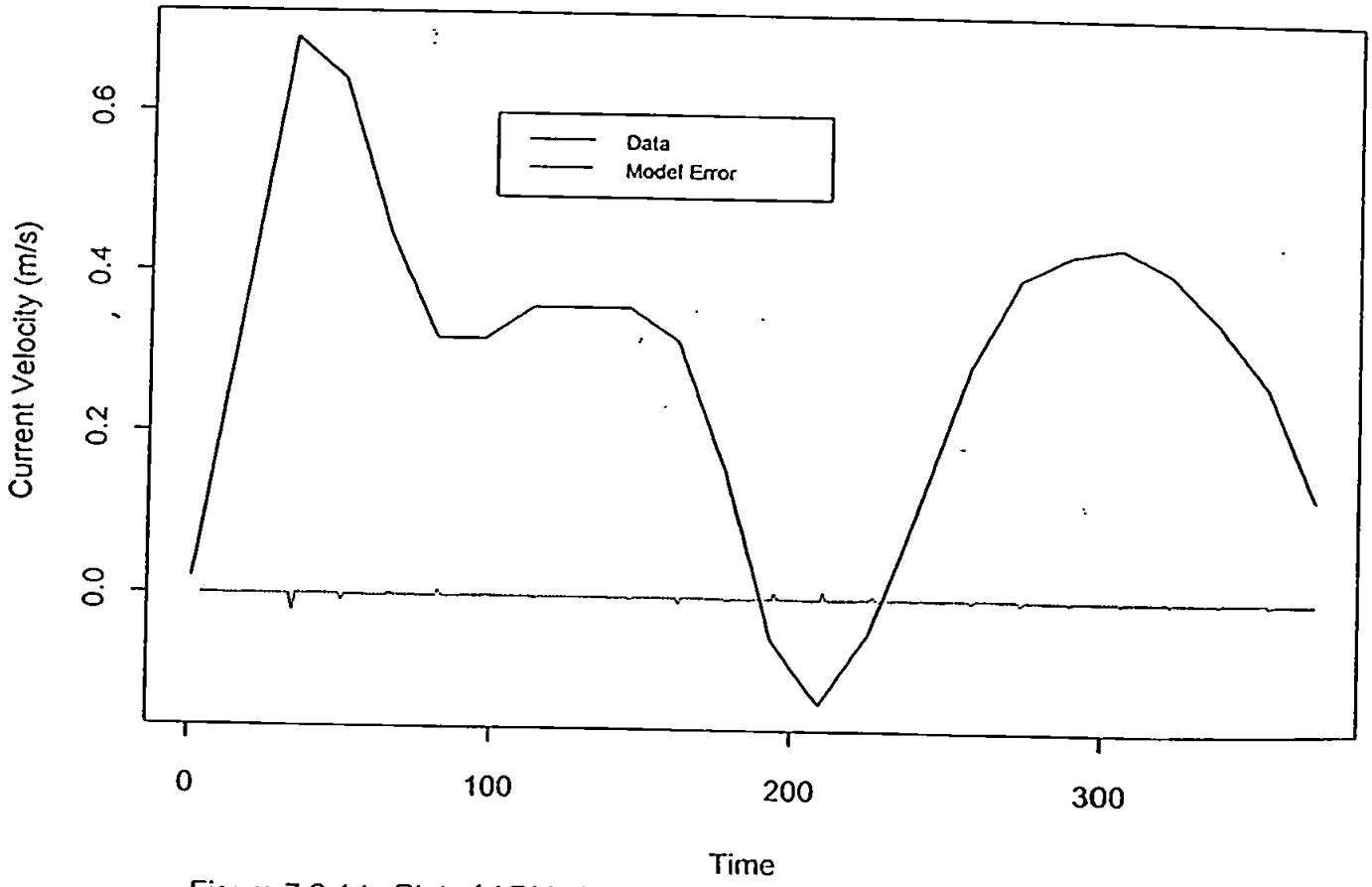


Figure 7.2.14. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.10H

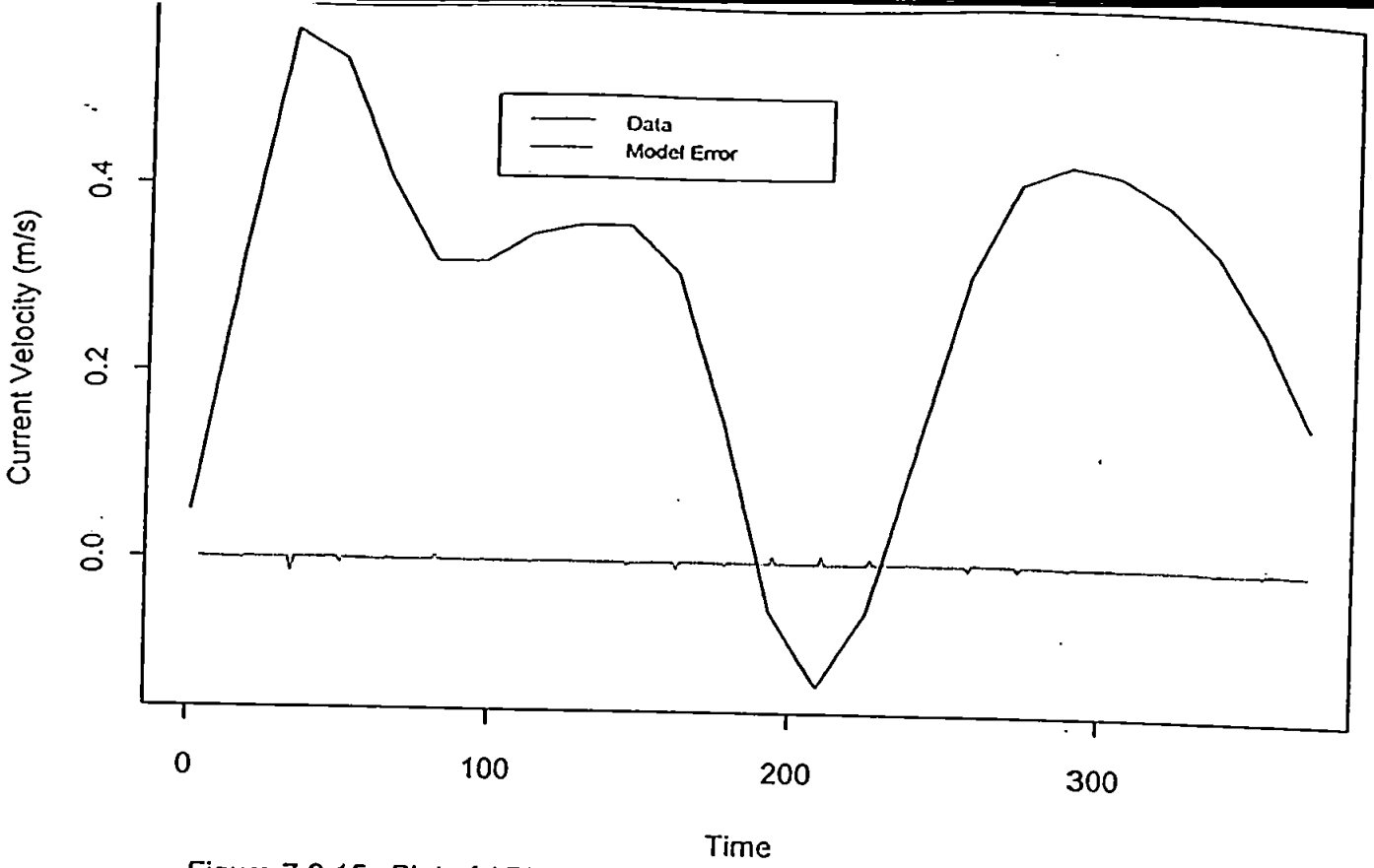


Figure 7.2.15. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.20H

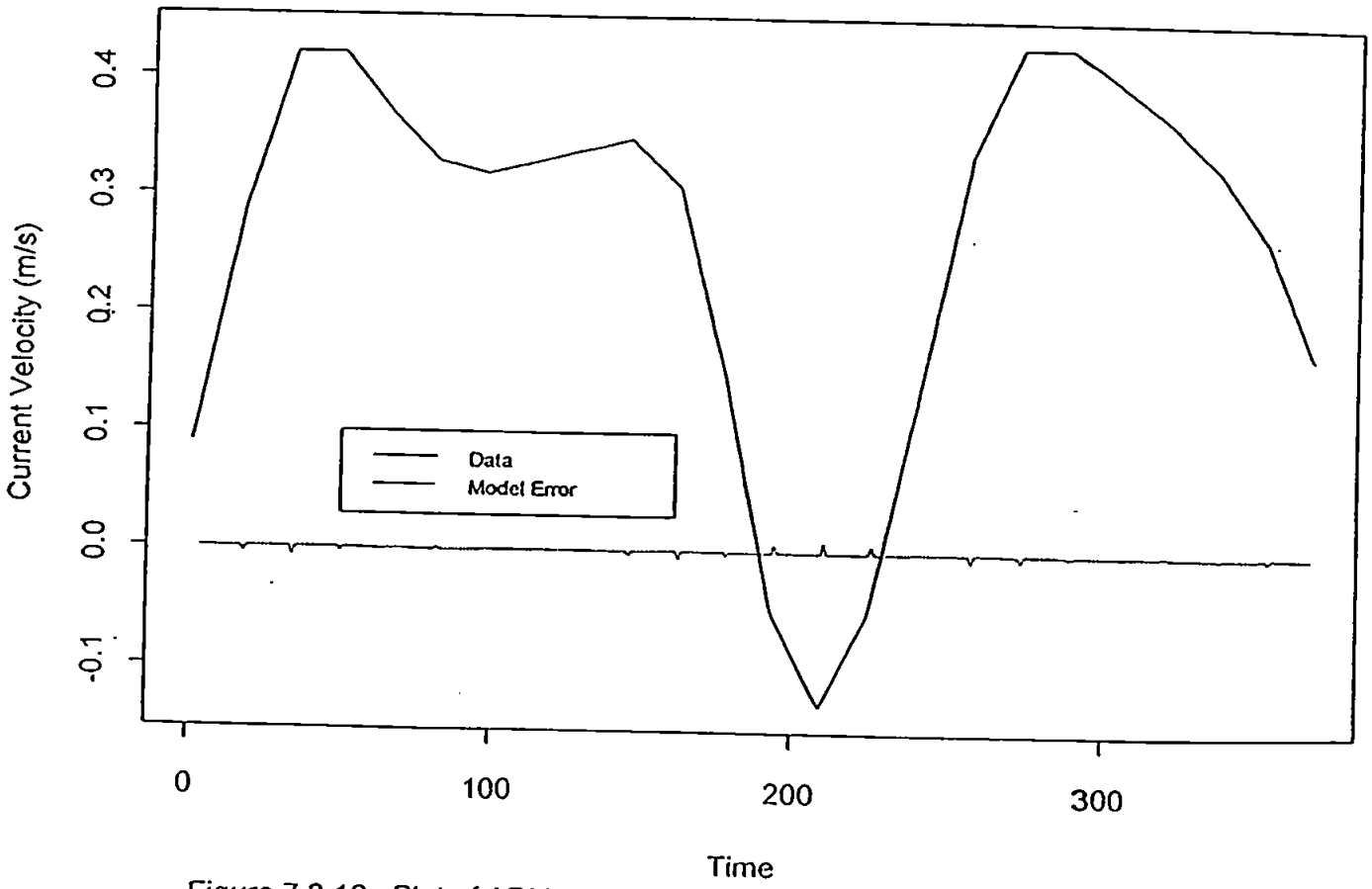


Figure 7.2.16. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.30H

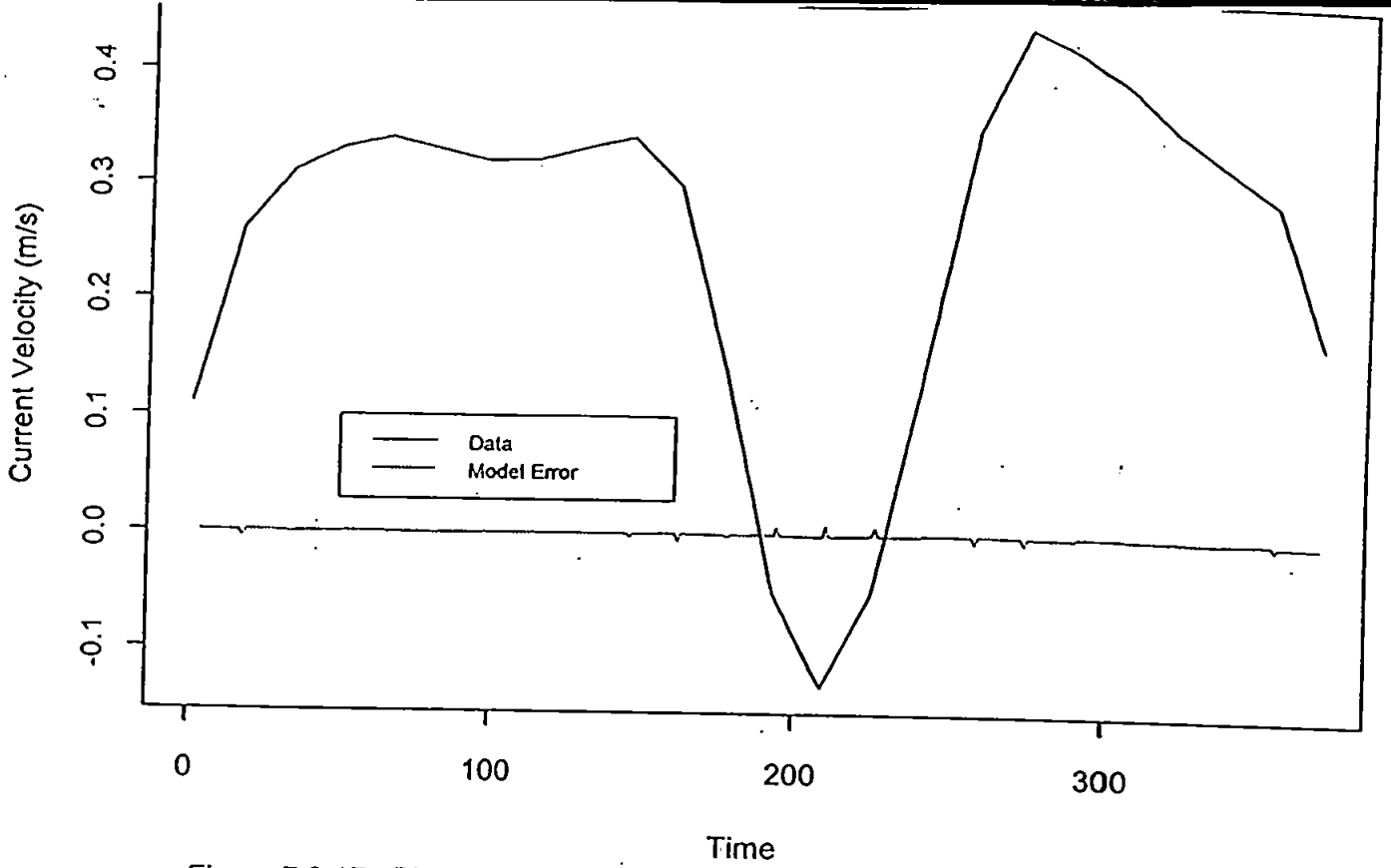


Figure 7.2.17. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.40H

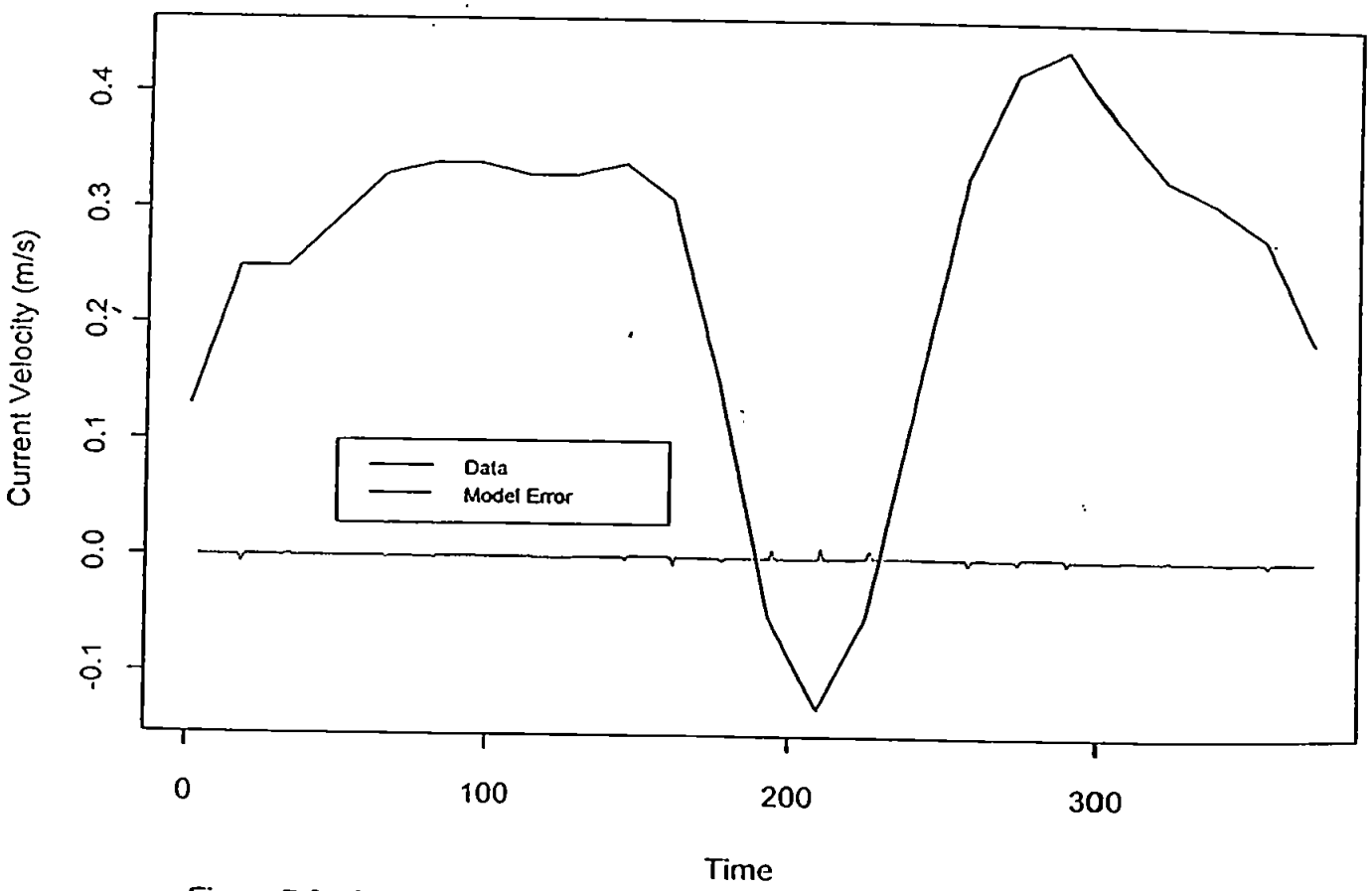


Figure 7.2.18. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.50H

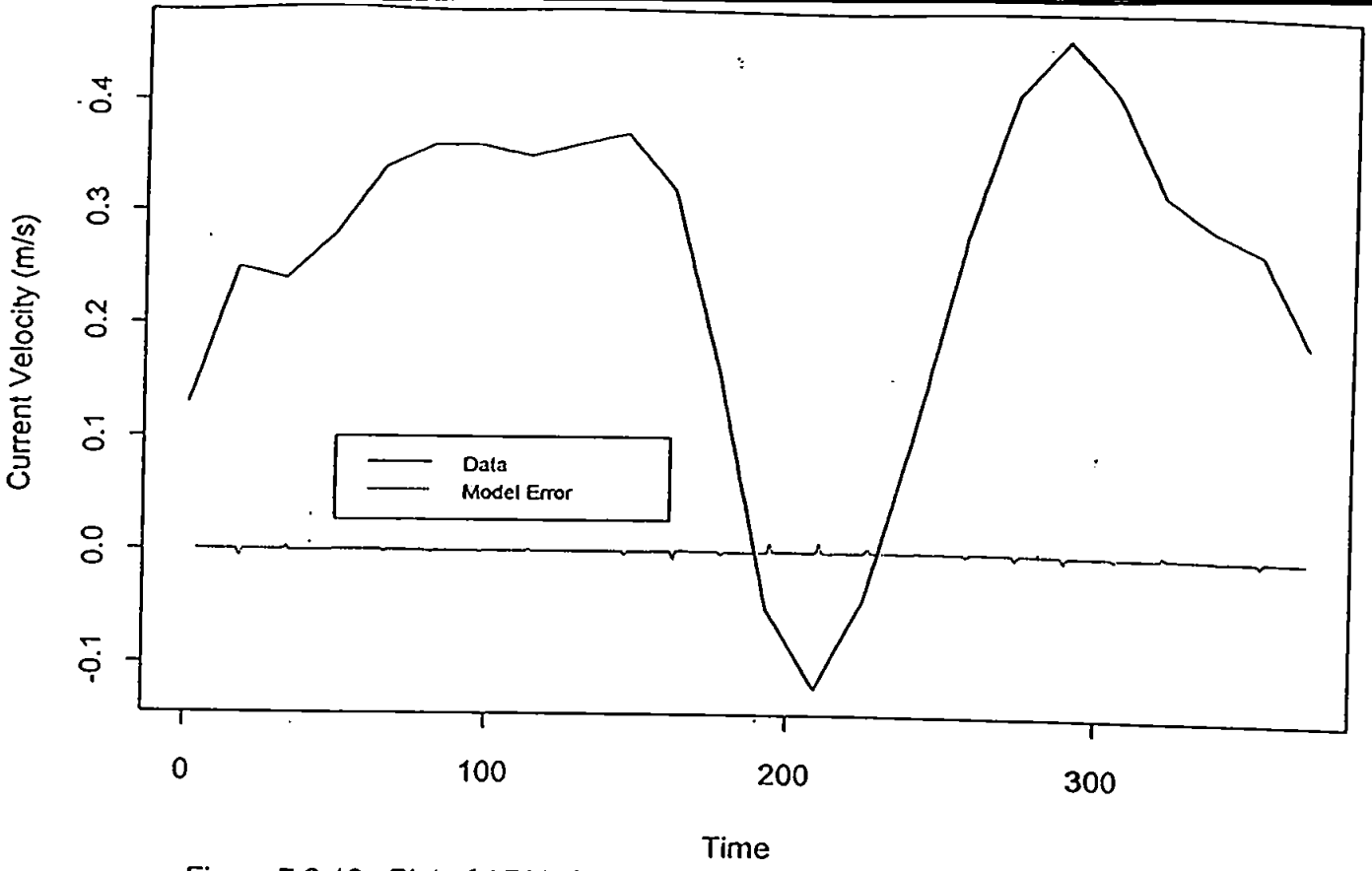


Figure 7.2.19. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.60H

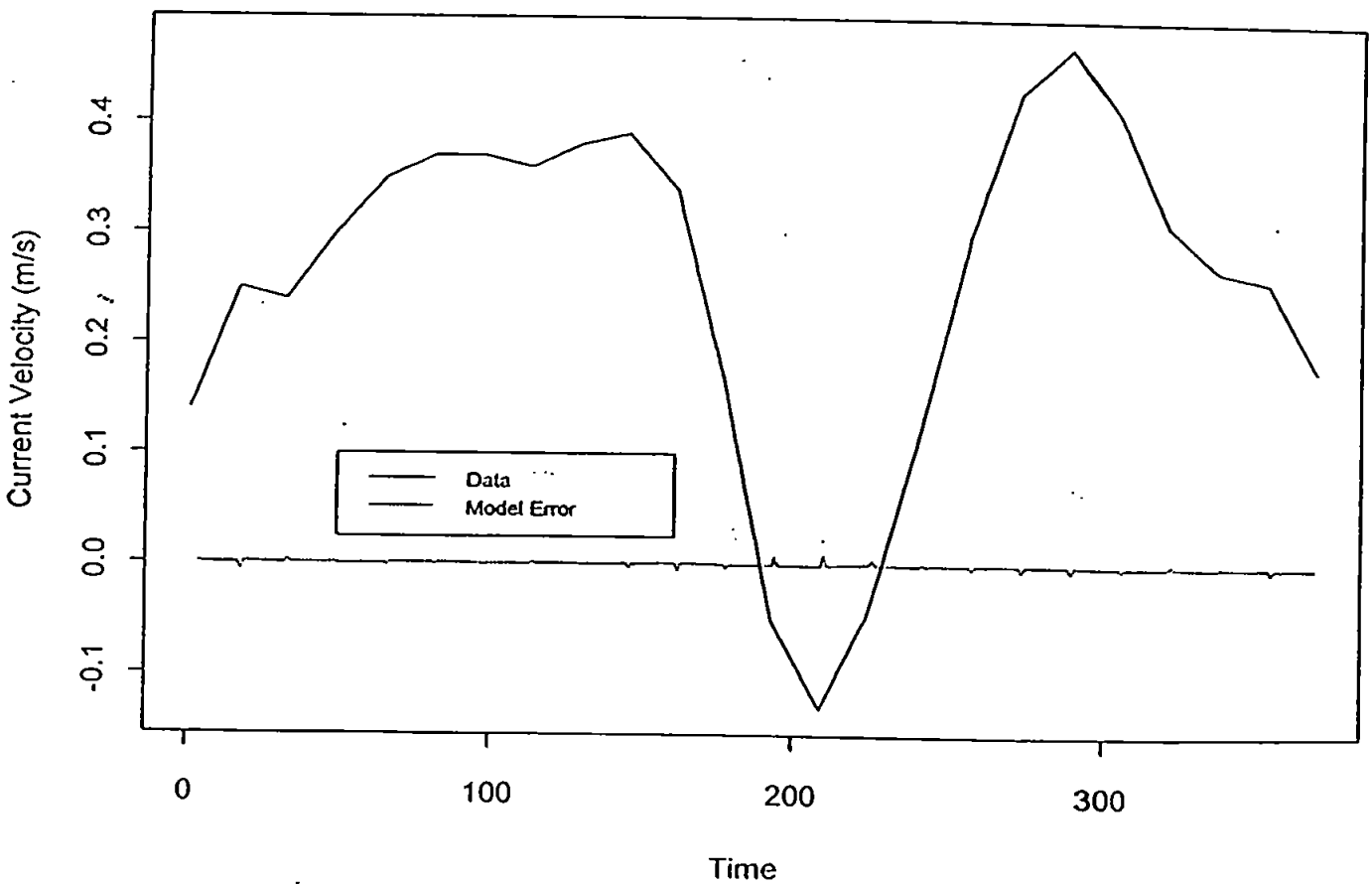


Figure 7.2.20. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.70H

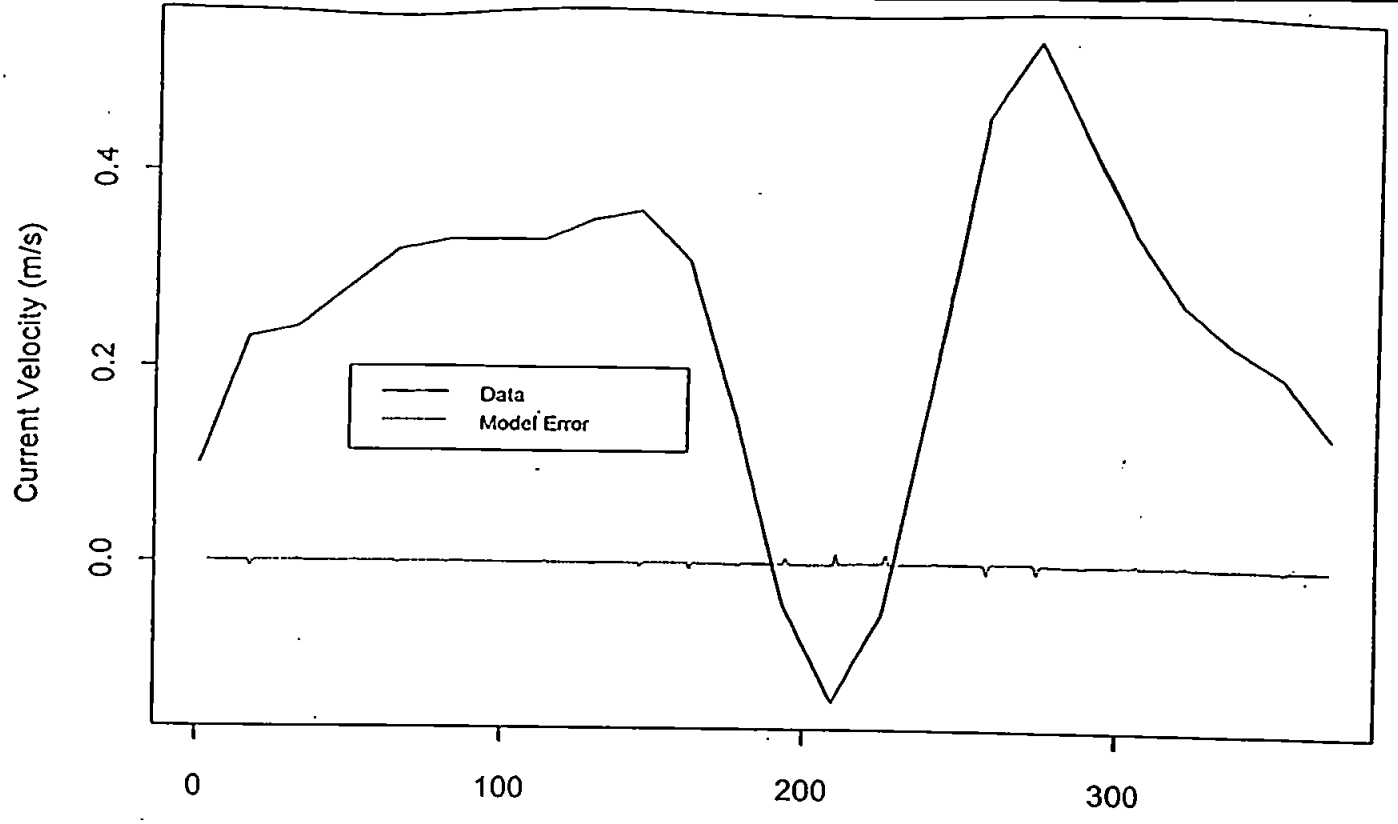


Figure 7.2.21. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.80H

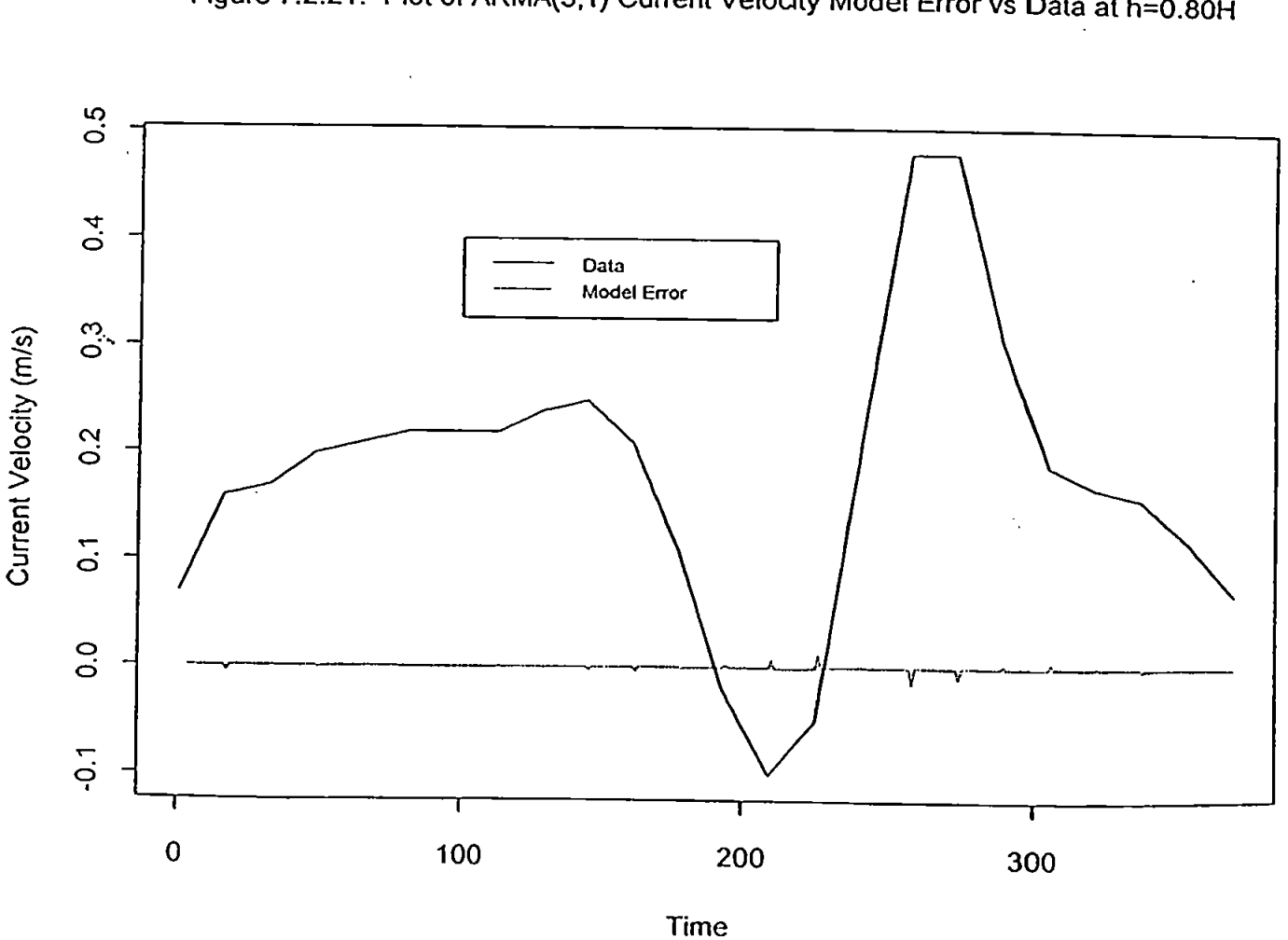


Figure 7.2.22. Plot of ARMA(3,1) Current Velocity Model Error vs Data at h=0.90H

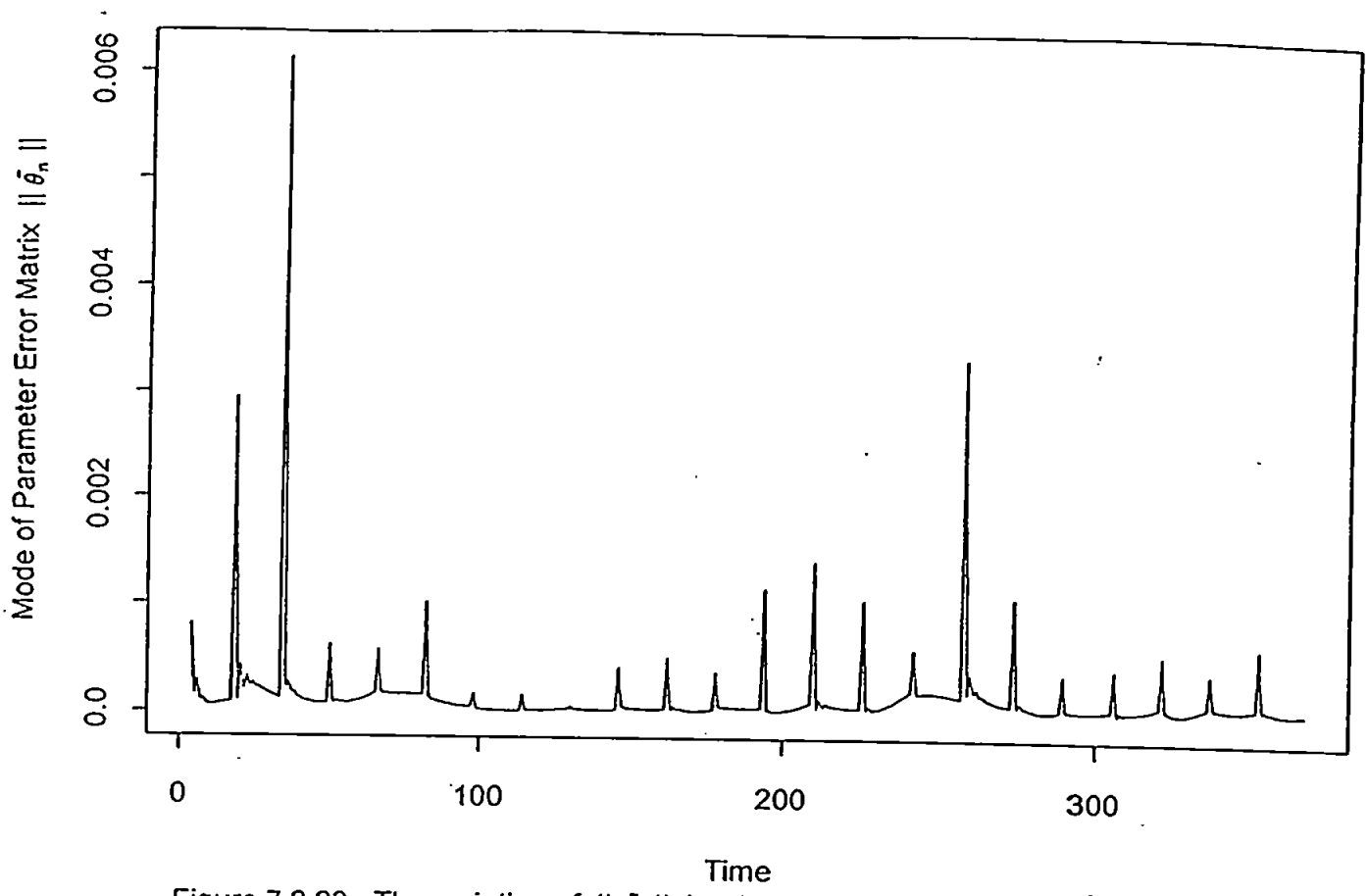


Figure 7.2.23. The variation of $\|\tilde{\theta}_n\|$ in ARMA(3,1) Current Velocity Model with time

(iv) Figures 7.2.13–7.2.22 show the simulation of the current velocity dynamics at different depth and Figure 7.2.23. gives the norm of the parameter matrix error dynamics in the current velocity model.
(v) The mode and the element of A_1, are comparatively large and the those of A_2, A_3 are comparatively small which shows that the more recent the time, the more effect is there on the current variation.
(vi) Since A_1 is strongly diagonally dominant, we can say that the larger the distance between given layers, the less effect is there on the layer variation.
(vii) The variation of the mode of parameter error matrix $\|\hat{\theta}_n\|$ with time is given in Figure 7.2.23. It is shown that the model parameter matrices are nearly constant matrices which means the good description of the model.

7.2.3 Model Comparison

In last two subsections, two kinds of multivariable time series current velocity model are presented. The model comparison between AR(5) and ARMA(3,1) is given in Table 7.2.5.

TABLE 7.2.5. The Model Comparison

Model	MPEE	MPVE	MPV	σ^2
AR(5)	0.0326627	0.0456293	0.0508491	2.05273e-05
ARMA(3,1)	0.028119	0.0393385	0.0061354	1.77705e-05

From Table 7.2.5, all the MPEE, MPVE, MPV and σ^2 of the ARMA model are smaller than those in AR model. It is shown that the MPV of ARMA are much improved which means that the parameter matrices of ARMA time series model are 'nearly constant matrices'. Therefore the multivariate model presented here is better description of the system, data fitting and prediction. This is because the ARMA model takes advantage of the information from the model error and estimation of system noise as well as under the assumption that the system noise is coloured noise. It is suggested that the ARMA Model is more suitable than AR Model in this special case.

7.3 Suspended Sediment Concentration Model

The resuspension, transport and deposition of suspended particular matter (SPM) play a crucial part in a range of marine processes, including benthic fluxes, biological productivity, biogeochemical cycling and pollutant dispersal. There are, however, surprisingly few data sets enabling detailed investigation of SPM dynamics because it has hitherto been difficult to monitor particle concentration, composition and behaviour over appropriate time and length scales. So it is very important to set up a proper SPM concentration model in the process of sediment transport.

In this section, two kinds of time series model are introduced. First, the univariate model similar to the current velocity model is used to describe the SPM concentration profile dynamics. Second, the multivariate model (ARMAX Model) is used to describe the SPM dynamics. The simulation and model comparison show that the latter one has advantage over the former one both in statistical analysis and geophysical illustration. Like the last section, the system identification technique is applied here to modelling the time series dynamics for SPM concentration by measuring depth profile current velocity and SPM concentration in situ, which provides a detail data set. The SPM time series model provides comprehensive modelling and prediction for SPM dynamics based on the data set.

7.3.1 AR Model and ARMA Model

In this subsection we focus on modelling the suspended sediment concentration dynamics. First of all, for simplicity, the univariate model for suspended sediment concentration profile is presented.

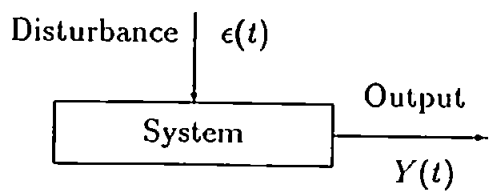


Fig 7.3.1 A Dynamic system with output $Y(t)$ and disturbance $\epsilon(t)$ where t denotes time.

Similar to the section 7.2, for simplicity, we first assume that the suspended sediment concentration variation $Y(t)$ are 'explained' solely in terms of its own past or by its time and surrounding value and uncontrollable disturbance $\epsilon(t)$, although of course the $Y(t)$ is also dependent on many other factors such as current velocity profile dynamics.

We assume that the time series model here that describes the suspended sediment concentration is a discrete multivariable time-invariant stochastic linear system and can be represented by the following AR(p) or ARMA(p, q) model like current velocity model presented in last section.

I. AR Model

1. Model Description

First, we assume that the suspended sediment profile model is a multivariable AR(p) model as follows:

$$A(z^{-1})Y_n = w_n \tag{7.23}$$

where w_n is the system noise and the restriction on it is presented in (3.66) and (3.67).

$$A(z^{-1}) = I_m + A_1 z^{-1} + ... + A_p z^{-p} \tag{7.24}$$

Y_n and w_n are m-dimensional vectors, z^{-1} is a unit delay operator and $A_i, (i = 1, ..., p)$ is $m \times m$ unknown matrices to be estimated. I_m is an $m \times m$ unit matrix.
Set

$$\theta^T = [-A_1, ..., -A_p]_{m \times d} \tag{7.25}$$

$$x_n^T = [Y_{n-1}^T, ..., Y_{n-p}^T]_{1 \times d} \tag{7.26}$$

$$d = m \times p \tag{7.27}$$

here θ_n is the estimate of θ at time n and $[.]_{m\times d}$ and $[.]_{1\times d}$ denote an $m \times d$ matrix and a $d-$dimensional row vector respectively.
It is easy to see that (7.23) also can be written as

$$Y_n = \theta^T x_n + w_n \tag{7.28}$$

We construct the vertical profile variable vector of suspended sediment concentration as the system output and it is a function of the water depth and time. Denote $Y_n(h)$ as the suspended sediment concentration at height h and time n (i.e. set water depth H, bottom $h = H$ and surface $h = 0$). So
$Y_n^T = [Y_n(h_1), Y_n(h_2), Y_n(h_3), ..., Y_n(h_m)]_{1\times m}$. where $0 \leq h_1 \leq h_2 \leq ... \leq h_m \leq H$.

In order to identify the system parameter matrix θ, we make use of the following recursive algorithms (3.55-3.57).

2. Order determination

The order determination technique is similar to the last section and the F-test results for our model candidates according to (4.11) and (4.12) are given as following in Table 7.3.1:

TABLE 7.3.1. The order comparison of suspended sediment concentration model

	MPVE	MPEE	MPV	σ^2
p=4	0.00858969	0.0209755	0.0415207	1.66488e-05
p=5	0.00947416	0.0221978	0.0249716	8.22566e-06
p=6	0.00948561	0.0222891	0.0322295	8.43719e-06

(i) Let AR(4)=$\mathcal{U}_1$, AR(5)=$\mathcal{U}_2$
$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{1.66488e-05-8.22566e-06}{8.22566e-06} = 373.763$ reject AR(4)

(ii) Let AR(5)=$\mathcal{U}_1$, AR(6)=$\mathcal{U}_2$
$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{8.22566e-06-8.43719e-06}{8.43719e-06} = -9.15097$ choose AR(5)

3. Simulation

According to Table 7.3.1, we choose a multivariable time series model which is a simpler form of (7.23) as following an AR(5) model:

$$Y_n = AY_{n-1} + BY_{n-2} + CY_{n-3} + DY_{n-4} + EY_{n-5} + w_n. \tag{7.29}$$

Set :

$$\theta^T = [A, B, C, D, E] \tag{7.30}$$

$$x_n^T = [Y_{n-1}, Y_{n-2}, Y_{n-3}, Y_{n-4}, Y_{n-5}]. \tag{7.31}$$

For solving equation (7.29), the ELSM (3.55)-(3.57) is applied (substitute Y_n for y_n) in (3.55). From (7.27), we know $p = 5, m = 10, d = 50$ and the time scale in here is $3\frac{3}{4}$ minutes per run. Since the initial value are need, the first fifth data as our initial value we really start the model at n=6.

The computation procedure of ELSM is as follows:

(i) Construct x_n according to (7.31) ($n \geq 5$).
(ii) Select initial values of θ_5 and R_5.
(iii) Calculate K_n, R_n and θ_n according to (3.55)-(3.57) based on the K_{n-1}, R_{n-1} , θ_{n-1} and x_n ($n \geq 5$).

The simulation results are given as follows:
(i) The five parameter matrices are:

$$A = \begin{pmatrix} 1.542 & 0.237 & 0.297 & 0.008 & -0.084 & -0.104 & 0.048 & -0.047 & -0.105 & 0.061 \\ 0.277 & 1.250 & 0.317 & 0.093 & 0.045 & -0.064 & 0.010 & 0.014 & -0.051 & 0.022 \\ 0.173 & 0.147 & 1.185 & 0.225 & 0.074 & 0.095 & 0.145 & -0.058 & -0.089 & 0.039 \\ 0.056 & 0.047 & 0.249 & 1.087 & 0.205 & 0.211 & 0.163 & -0.043 & -0.041 & 0.016 \\ 0.004 & 0.017 & 0.075 & 0.225 & 1.252 & 0.179 & 0.215 & -0.018 & -0.001 & 0.021 \\ -0.010 & -0.065 & 0.076 & 0.209 & 0.152 & 1.271 & 0.285 & 0.045 & -0.066 & 0.056 \\ -0.002 & -0.063 & 0.037 & 0.082 & 0.053 & 0.186 & 1.338 & 0.279 & 0.038 & -0.035 \\ -0.017 & 0.008 & 0.001 & -0.006 & 0.007 & 0.080 & 0.423 & 1.229 & 0.093 & 0.086 \\ -0.047 & -0.005 & 0.003 & 0.069 & 0.114 & -0.014 & 0.135 & -0.021 & 1.232 & 0.428 \\ -0.008 & -0.046 & 0.002 & 0.088 & 0.064 & 0.013 & -0.105 & -0.178 & 0.257 & 1.819 \end{pmatrix}$$

$$B = \begin{pmatrix}
-0.406 & -0.404 & -0.258 & -0.074 & 0.056 & 0.071 & 0.083 & 0.086 & 0.074 & -0.056 \\
-0.428 & -0.074 & -0.269 & -0.138 & -0.043 & 0.034 & 0.016 & -0.064 & -0.005 & 0.052 \\
-0.257 & -0.281 & 0.145 & -0.217 & -0.105 & -0.129 & -0.085 & -0.087 & -0.011 & 0.057 \\
-0.084 & -0.174 & -0.204 & 0.182 & -0.240 & -0.224 & -0.149 & -0.111 & -0.055 & -0.063 \\
0.002 & -0.102 & -0.093 & -0.218 & 0.041 & -0.230 & -0.200 & -0.175 & -0.098 & 0.061 \\
0.045 & -0.021 & -0.087 & -0.198 & -0.176 & 0.027 & -0.306 & -0.282 & -0.053 & 0.072 \\
0.046 & -0.003 & -0.006 & -0.051 & -0.118 & -0.258 & -0.006 & -0.522 & -0.135 & 0.167 \\
0.044 & -0.029 & 0.015 & 0.020 & -0.029 & -0.162 & -0.400 & -0.138 & -0.200 & -0.014 \\
0.061 & 0.027 & -0.025 & -0.089 & -0.123 & -0.053 & -0.010 & -0.097 & -0.090 & -0.478 \\
-0.002 & 0.116 & -0.054 & -0.108 & -0.082 & -0.048 & 0.263 & 0.184 & -0.500 & -0.659
\end{pmatrix}$$

$$C = \begin{pmatrix}
-0.110 & 0.090 & 0.039 & 0.044 & 0.028 & -0.045 & -0.055 & -0.035 & 0.060 & -0.046 \\
0.034 & -0.015 & -0.039 & -0.010 & -0.008 & 0.037 & 0.043 & -0.037 & 0.018 & -0.030 \\
0.014 & 0.004 & 0.022 & -0.064 & 0.028 & -0.030 & -0.021 & 0.010 & 0.068 & -0.002 \\
0.028 & 0.012 & -0.054 & 0.049 & -0.046 & -0.042 & -0.017 & 0.030 & 0.057 & 0.033 \\
0.010 & 0.008 & 0.040 & -0.037 & -0.071 & 0.001 & -0.029 & 0.036 & 0.056 & 0.035 \\
-0.008 & 0.049 & -0.006 & -0.059 & 0.031 & -0.063 & -0.035 & 0.050 & 0.102 & -0.032 \\
-0.038 & 0.036 & -0.012 & 0.009 & 0.043 & -0.004 & 0.004 & -0.064 & 0.012 & -0.001 \\
-0.022 & -0.020 & -0.024 & 0.049 & 0.080 & 0.056 & -0.097 & -0.029 & -0.002 & -0.039 \\
0.005 & -0.022 & -0.011 & 0.002 & 0.013 & 0.071 & -0.024 & 0.001 & -0.060 & -0.002 \\
-0.015 & 0.009 & -0.018 & -0.007 & 0.060 & -0.028 & 0.015 & 0.009 & 0.029 & -0.096
\end{pmatrix}$$

$$D = \begin{pmatrix}
-0.057 & 0.084 & 0.041 & 0.020 & -0.014 & -0.030 & -0.001 & -0.063 & 0.030 & 0.022 \\
0.068 & -0.067 & 0.008 & 0.013 & 0.006 & -0.007 & 0.032 & -0.016 & 0.029 & -0.027 \\
0.026 & 0.065 & -0.119 & 0.007 & 0.030 & 0.000 & 0.011 & 0.023 & 0.039 & -0.049 \\
-0.007 & 0.053 & 0.001 & -0.101 & 0.038 & 0.016 & 0.024 & 0.029 & 0.022 & -0.057 \\
-0.022 & 0.045 & 0.010 & 0.010 & -0.092 & 0.021 & 0.015 & 0.065 & 0.036 & -0.077 \\
-0.024 & 0.045 & -0.001 & -0.001 & 0.033 & -0.124 & 0.043 & 0.103 & 0.036 & -0.102 \\
-0.017 & 0.040 & -0.033 & -0.025 & 0.031 & 0.020 & -0.091 & 0.125 & 0.041 & -0.095 \\
-0.003 & 0.032 & -0.041 & -0.024 & 0.014 & 0.018 & 0.038 & -0.015 & 0.052 & -0.059 \\
-0.020 & 0.008 & 0.000 & 0.014 & 0.036 & 0.009 & -0.034 & 0.011 & -0.065 & 0.062 \\
-0.009 & -0.032 & 0.023 & 0.034 & 0.055 & -0.005 & -0.082 & -0.089 & 0.121 & 0.025
\end{pmatrix}$$

$$E = \begin{pmatrix} 0.013 & -0.035 & -0.009 & -0.036 & -0.054 & 0.070 & 0.133 & -0.067 & -0.124 & 0.078 \\ 0.064 & -0.137 & 0.037 & 0.010 & -0.008 & -0.041 & 0.043 & 0.027 & -0.046 & 0.025 \\ 0.048 & 0.078 & -0.254 & 0.031 & -0.010 & 0.004 & 0.079 & 0.056 & -0.033 & -0.030 \\ 0.010 & 0.076 & 0.012 & -0.271 & 0.072 & -0.015 & 0.082 & 0.059 & -0.002 & -0.048 \\ 0.004 & 0.069 & -0.075 & 0.008 & -0.101 & -0.022 & 0.057 & 0.070 & 0.026 & -0.056 \\ -0.003 & 0.040 & -0.056 & 0.045 & 0.028 & -0.192 & 0.066 & 0.077 & -0.048 & -0.011 \\ 0.007 & 0.041 & -0.051 & -0.021 & 0.043 & -0.001 & -0.217 & 0.178 & 0.072 & -0.059 \\ -0.017 & 0.087 & -0.032 & -0.050 & -0.017 & -0.064 & 0.113 & -0.085 & 0.078 & -0.018 \\ -0.018 & 0.082 & -0.038 & -0.035 & 0.009 & -0.098 & 0.056 & 0.086 & -0.046 & -0.012 \\ 0.007 & 0.034 & -0.006 & -0.024 & -0.082 & -0.001 & 0.033 & 0.027 & 0.088 & -0.094 \end{pmatrix}$$

ii)

TABLE 7.3.2. The mode of the parameter matrices in AR(5) suspended sediment concentration model

	A	B	C	D	E
p=5	2.004459	1.051542	0.266854	0.285625	0.398206

(iii)

MPEE=0.00947416 mg/l

MPVE=0.00221978 mg/l

MPV= 0.0249716

σ^2=8.22566 e-06

(iv) Figures 7.3.2–7.3.11 show the simulation of the suspended sediment concentration dynamics at different depth and Figure 7.3.12. gives the norm of the parameter matrix error dynamics in the AR(5) SPM concentration model.

(v) The mode and the element of A, B are comparatively large and those of C, D, E are comparatively small which shows that the more recent the time, the more effect is there on the current variation.

(vi) Since A is strongly diagonally dominant, we can say that the larger the distance between given layers, the less effect is there on the layer variation.

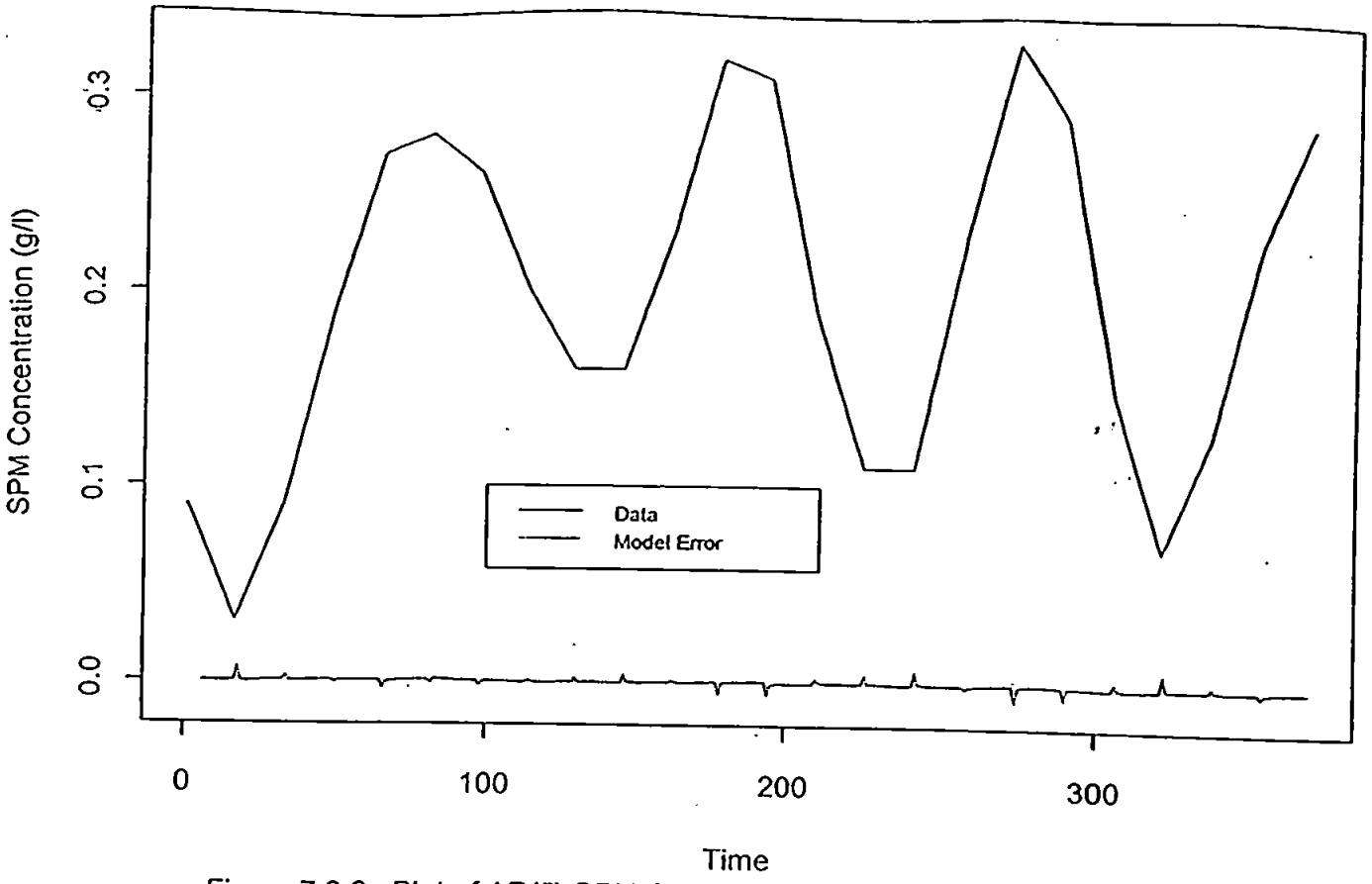


Figure 7.3.2. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.05H

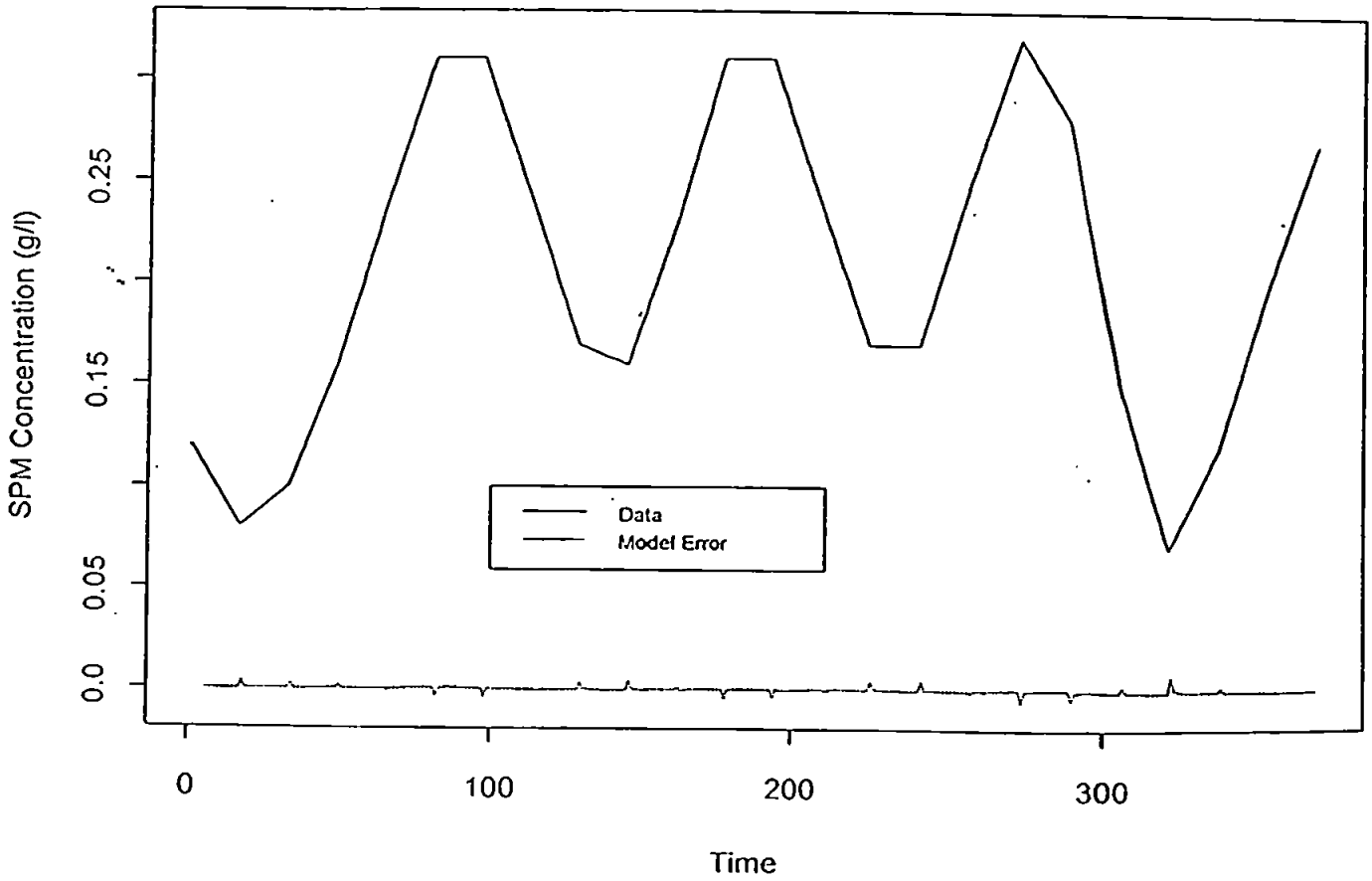


Figure 7.3.3. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.10H

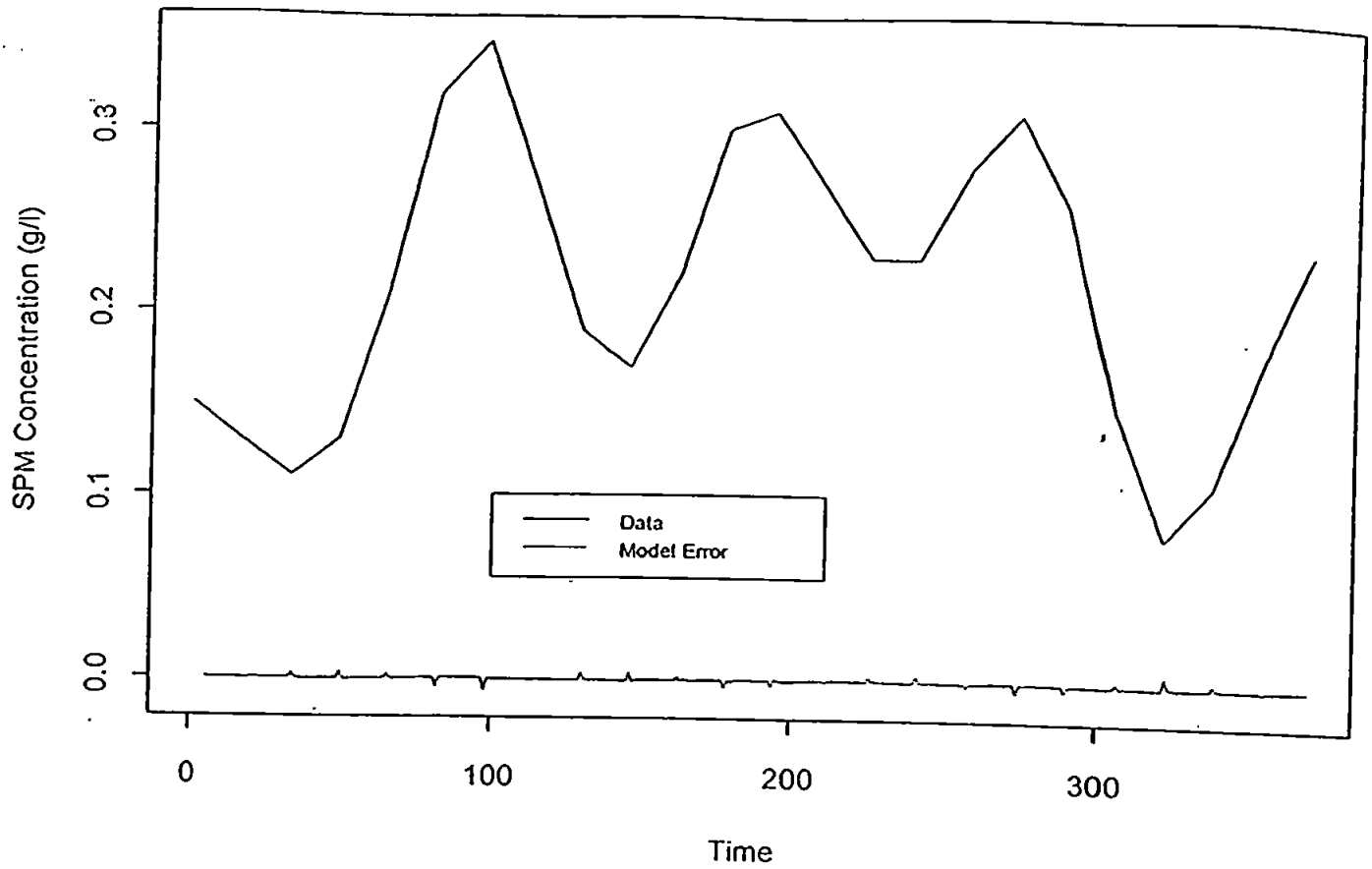


Figure 7.3.4. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.20H.

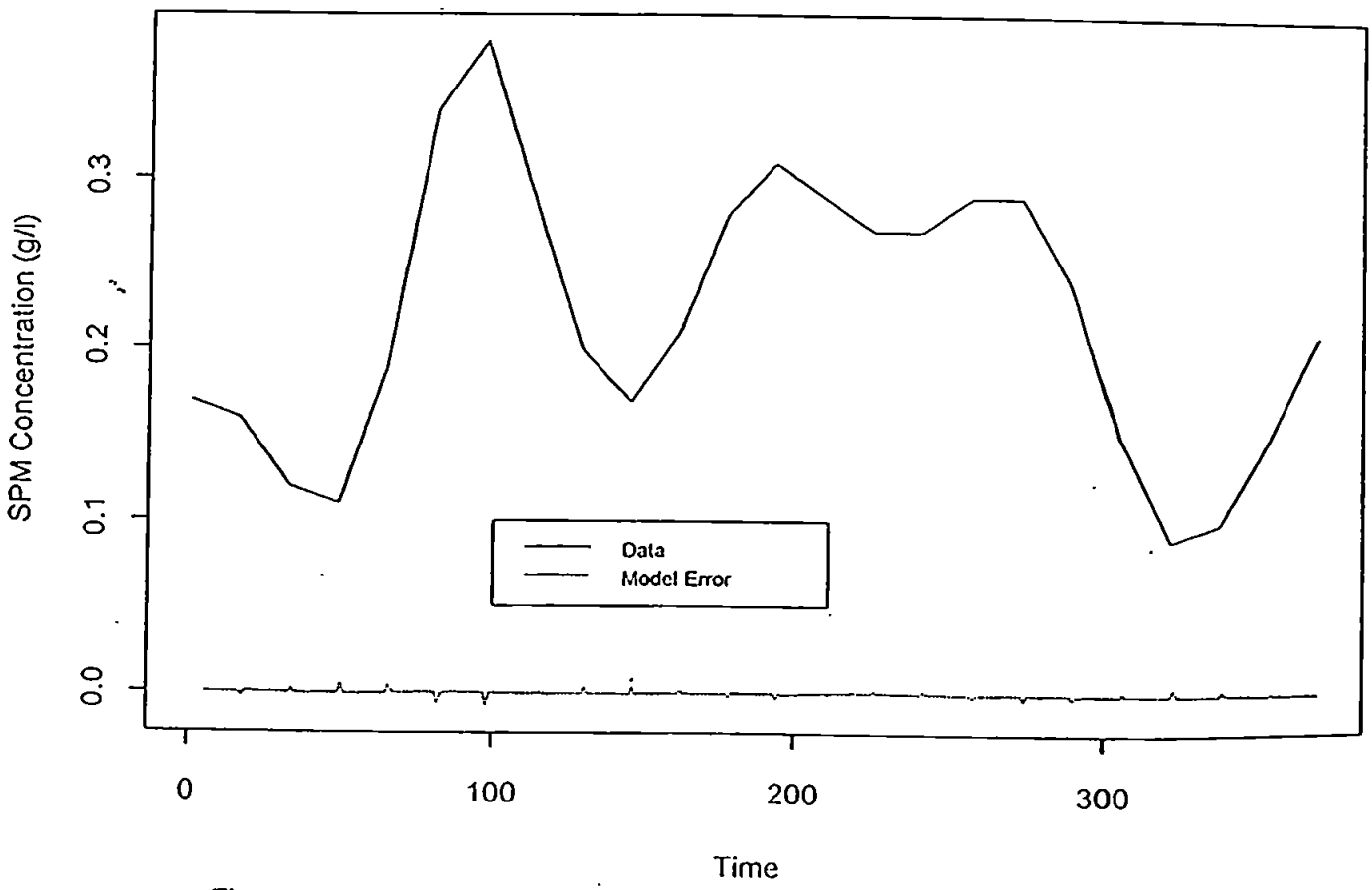


Figure 7.3.5. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.30H

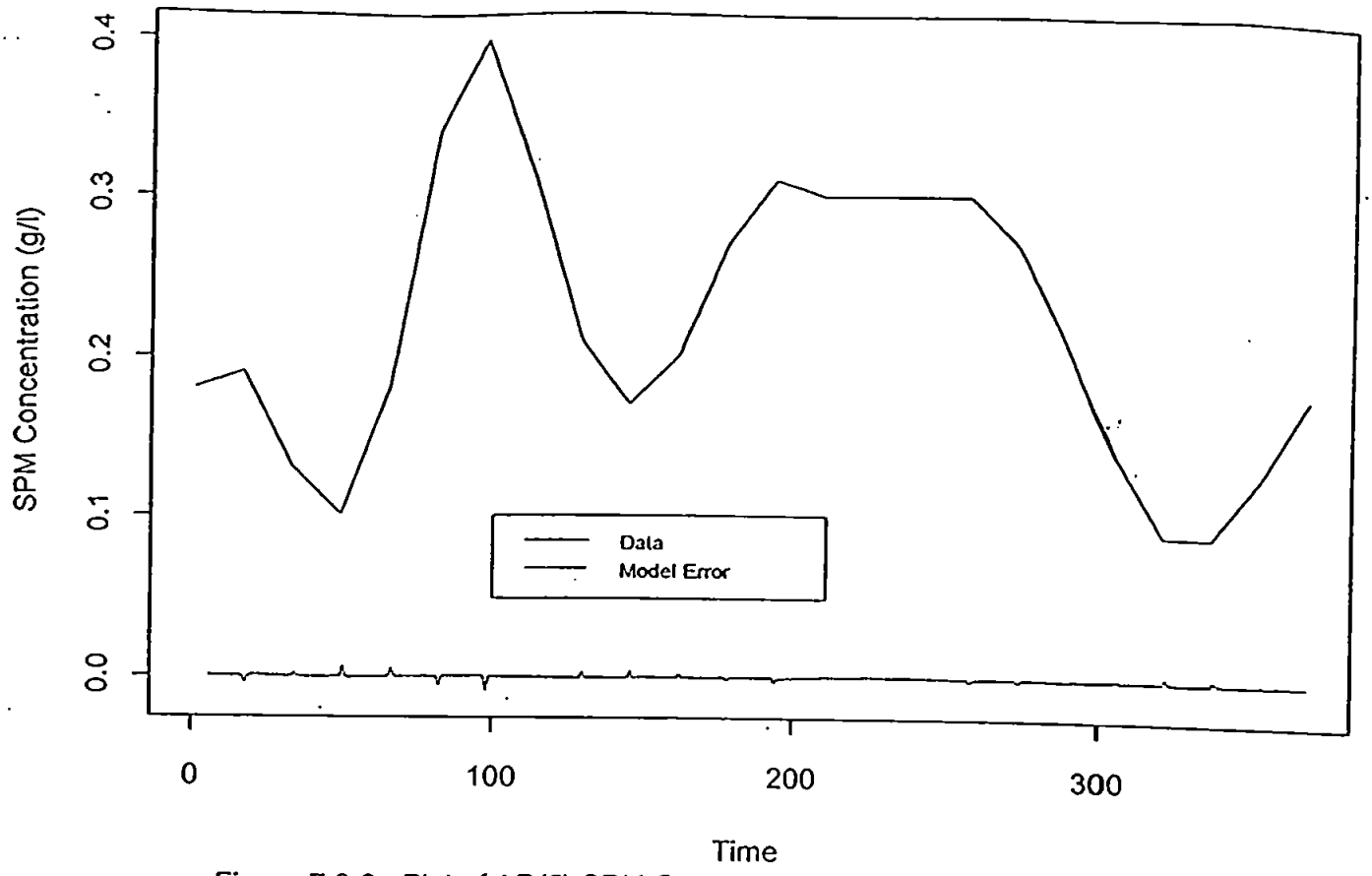


Figure 7.3.6. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.40H

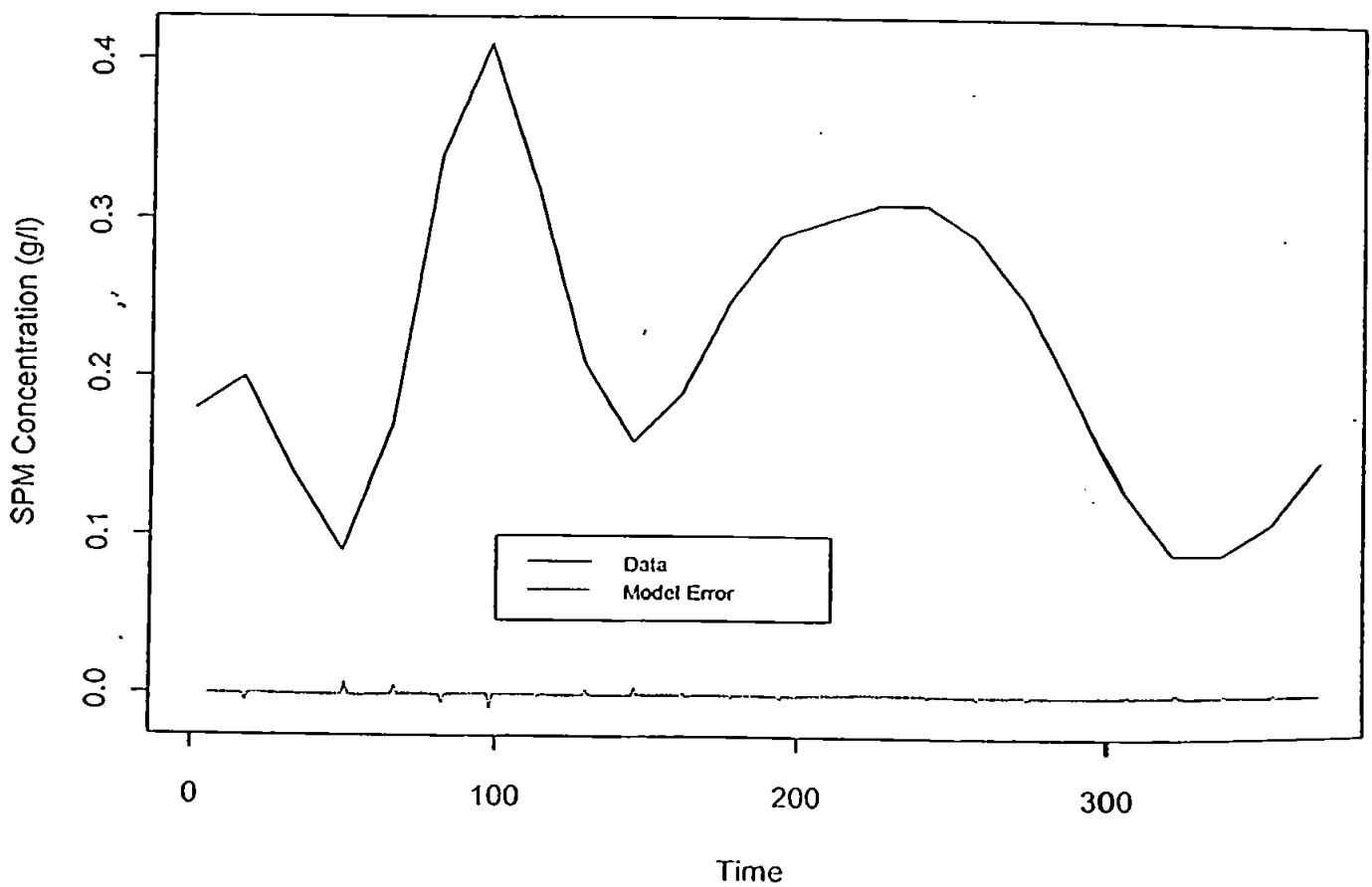


Figure 7.3.7. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.50H

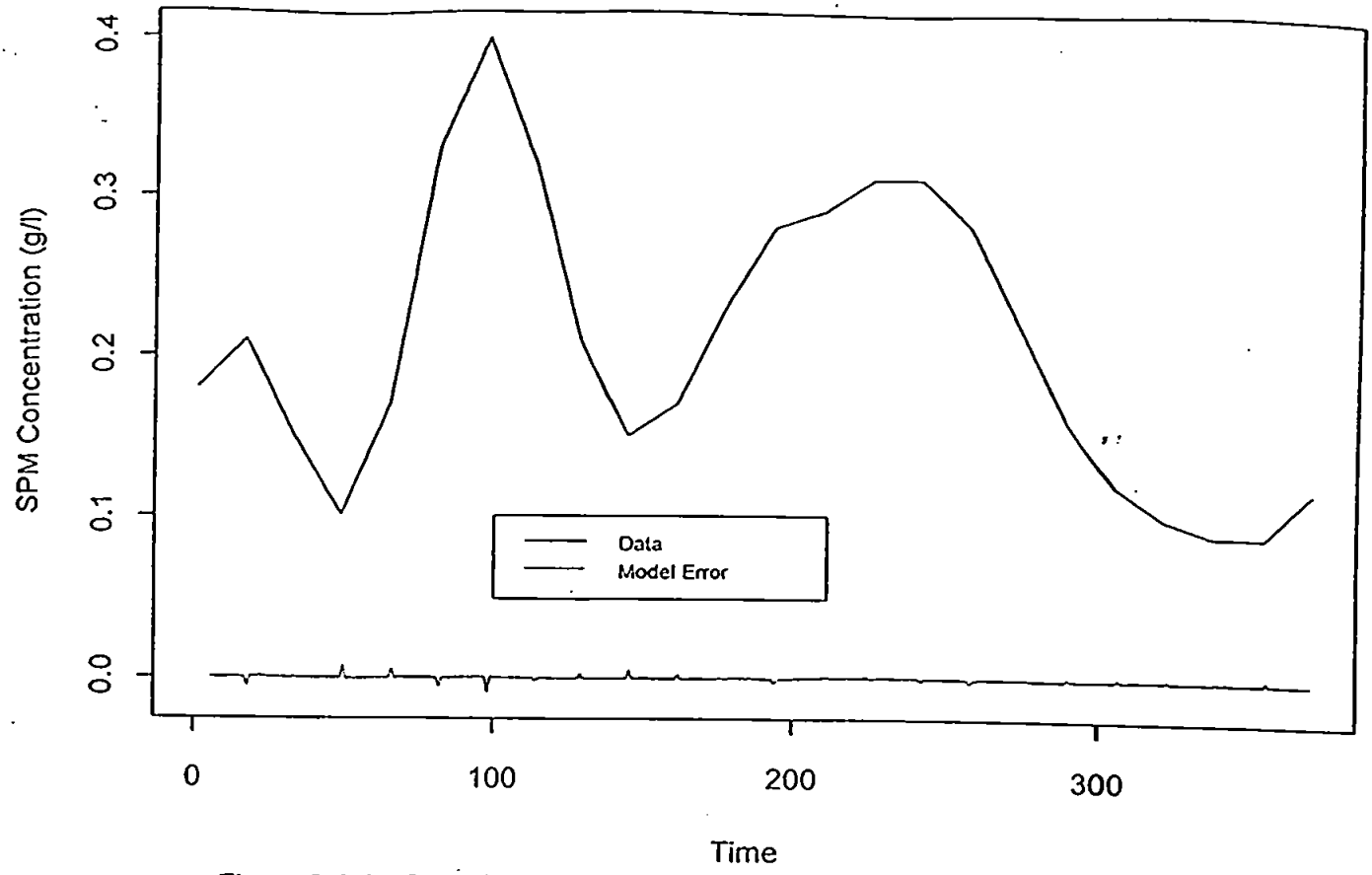


Figure 7.3.8. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.60H

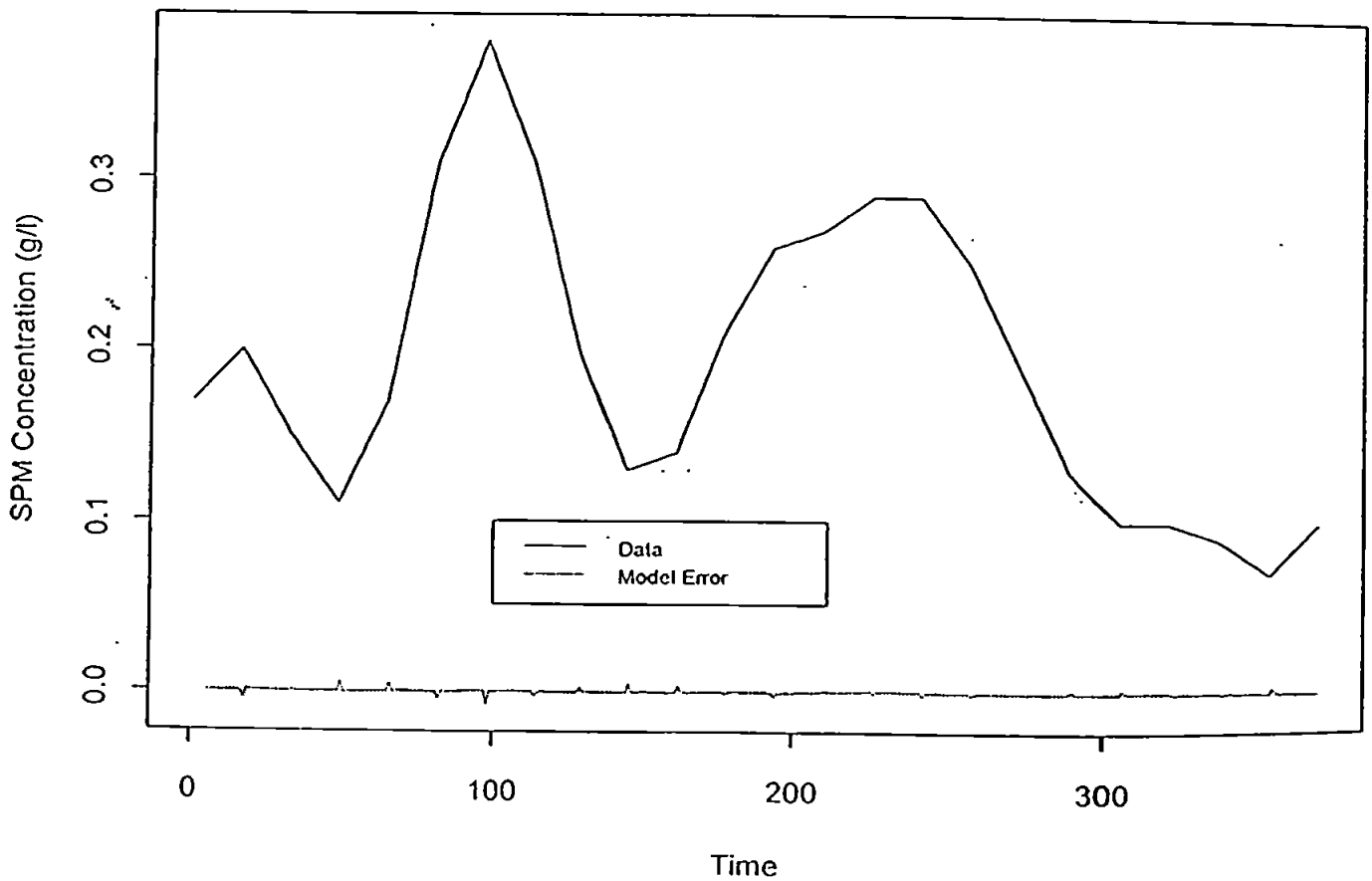


Figure 7.3.9. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.70H

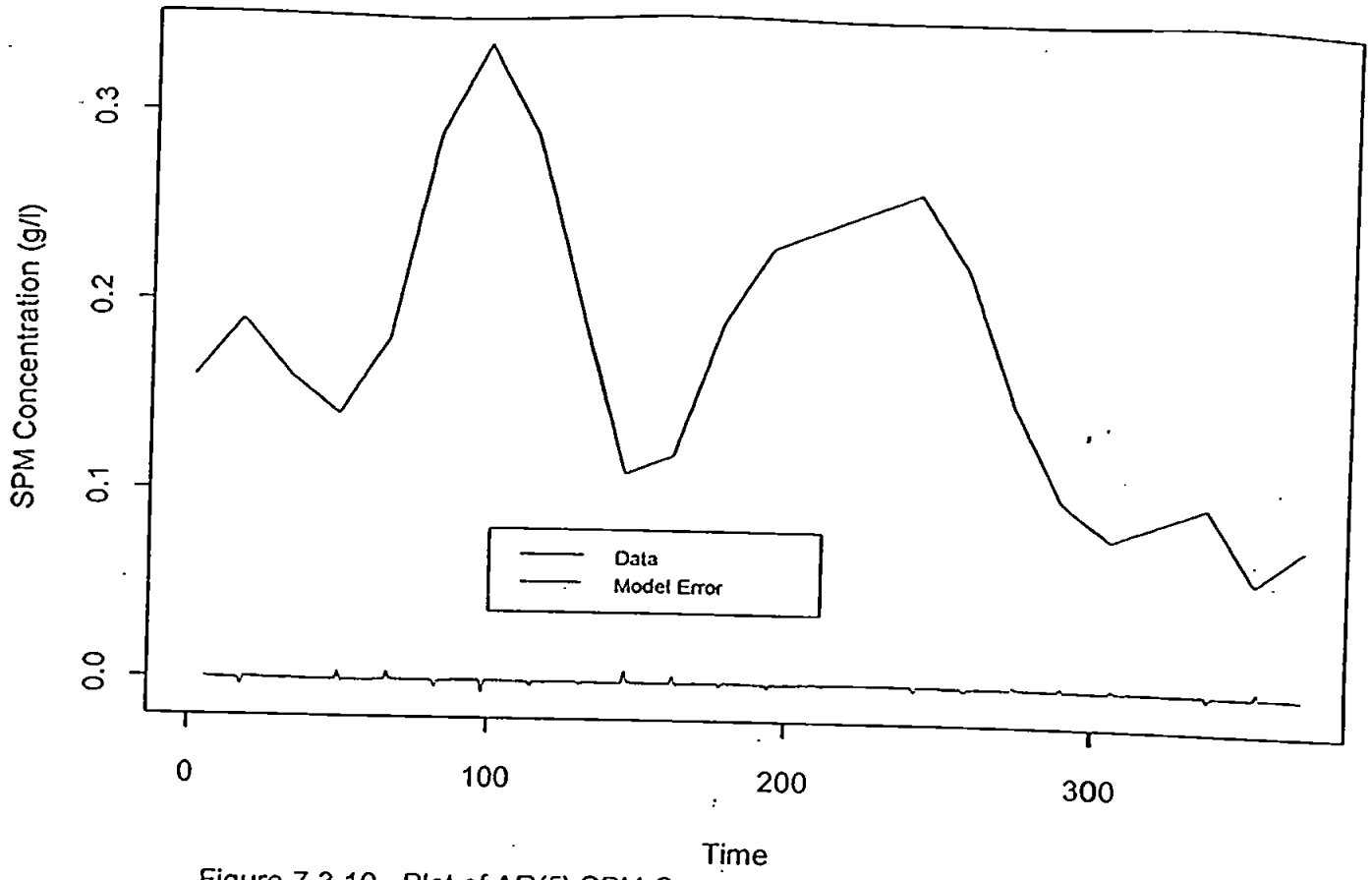


Figure 7.3.10. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.80H

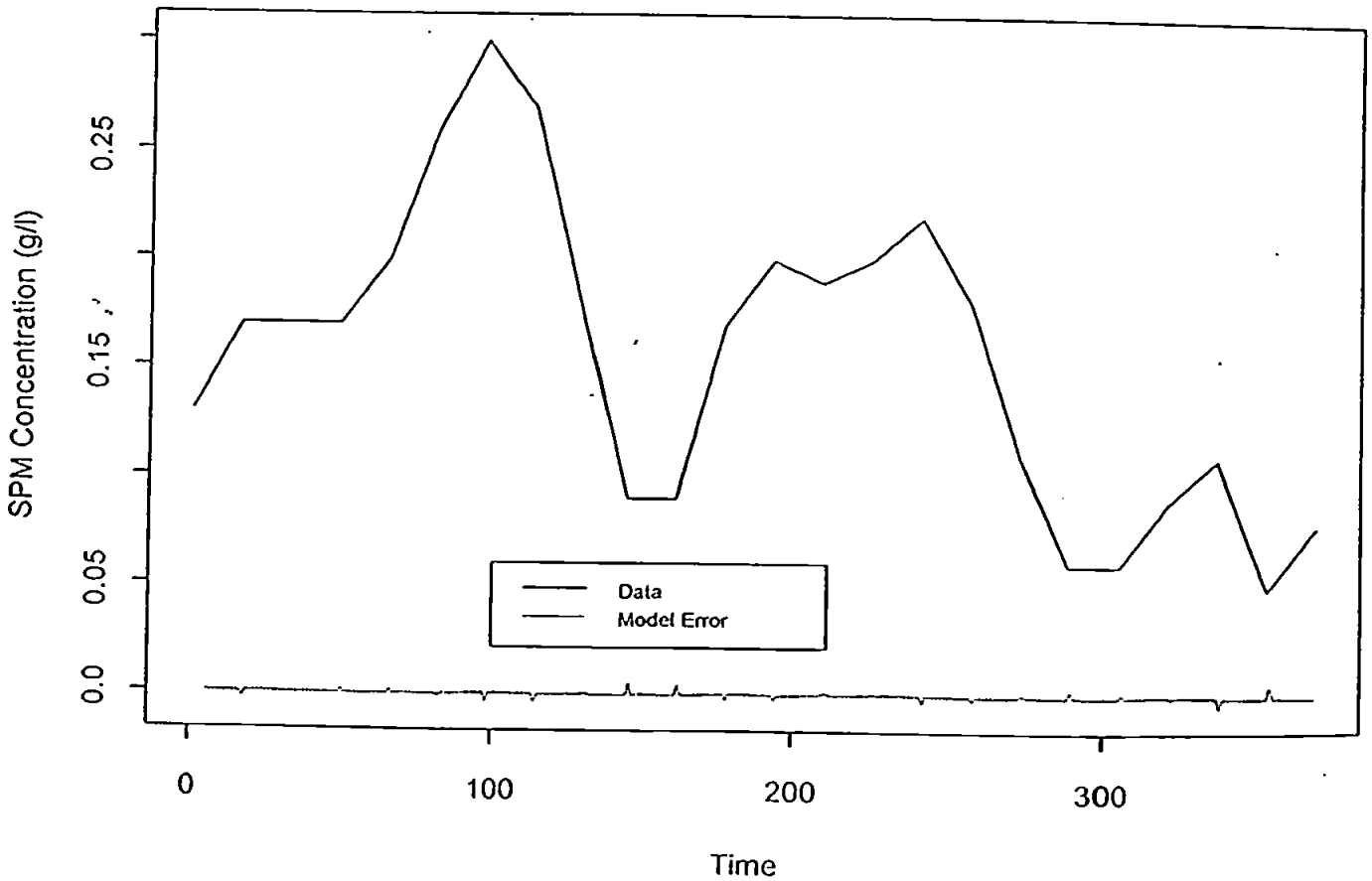


Figure 7.3.11. Plot of AR(5) SPM Concentration Model Error vs Data at h=0.90H

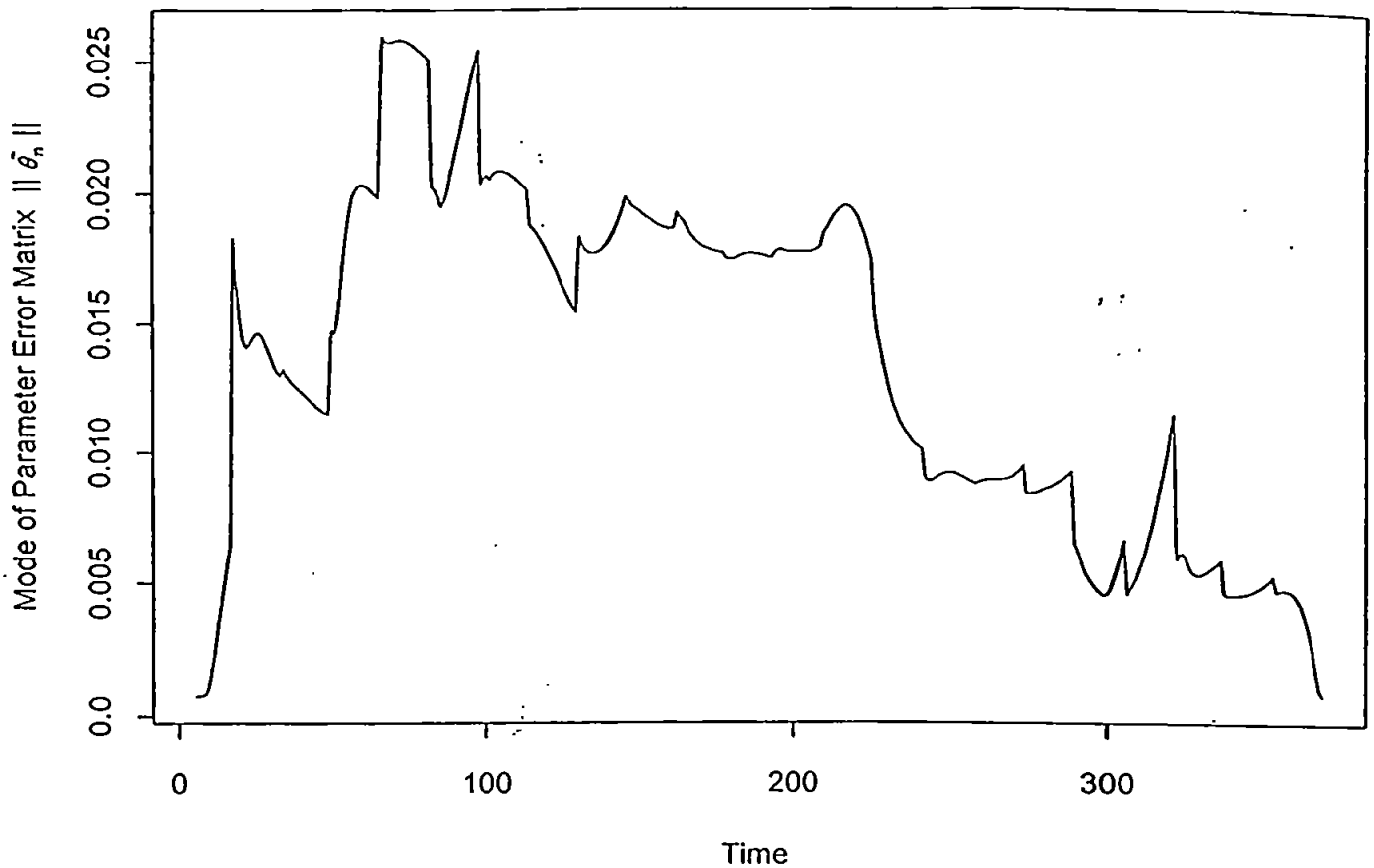


Figure 7.3.12. The variation of || $\tilde{\theta}_n$ || in AR(5) SPM Concentration Model with time

II. ARMA Model

1. Model Description

Now, we generalise the model structure and assume the SPM concentration dynamic system is disturbed by correlated noise series that the suspended sediment profile model is a multivariable ARMA(p,q) model as follows:

$$A(z^{-1})Y_n = C(z^{-1})w_n \tag{7.32}$$

where w_n is the system noise and the restriction on it presented in (3.66) and (3.67).

$$A(z^{-1}) = I_m + A_1 z^{-1} + ... + A_p z^{-p}$$
$$C(z^{-1}) = I_m + C_1 z^{-1} + ... + C_q z^{-q}$$

Y_n and w_n are m-dimensional vectors, z^{-1} is a unit delay operator and $A_i, C_j (i = 1, ..., p; j = 1, ..., q)$ is $m \times m$ unknown matrices to be estimated. I_m is an $m \times m$ unit matrix.
Set

$$\theta^T = [-A_1, ..., -A_p, C_1, ..., C_q]_{m \times d} \tag{7.33}$$

$$x_n^T = [Y_{n-1}^T, ..., Y_{n-p}^T, e_{n-1}^T, ..., e_{n-q}^T]_{1 \times d} \tag{7.34}$$

$$d = m \times p \tag{7.35}$$

$$e_n = y_n - \theta_n^T x_n \tag{7.36}$$

here θ_n is the estimate of θ at time n and $[.]_{m \times d}$ and $[.]_{1 \times d}$ denote an $m \times d$ matrix and a $d-$dimensional row vector respectively.
It is easy to see that (7.32) can also be written as

$$Y_n = \theta^T x_n + C(z^{-1})w_n + e_n - C(z^{-1})e_n. \tag{7.37}$$

We construct the vertical profile variable vector of suspended sediment concentration as the system output and it is a function of the water depth and time as we mentioned before.

In order to identify the system parameter matrix θ, we make use of the recursive algorithms (3.55–3.57).

2. Order determination

The F-test results for our model candidates according to (4.11) and (4.12) are given as following in Table 7.3.3:

TABLE 7.3.3 The order comparison of suspended sediment concentration model

	MPVE	MPEE	MPV	σ^2
p=2, q=1	0.00871258	0.0209343	0.00236536	1.06983e-05
p=3, q=1	0.00883307	0.0209273	0.00255276	7.8781e-06
p=4, q=1	0.00881297	0.0207835	0.00263308	8.20009e-06

(i) Let ARMA(2,1)=$\mathcal{U}_1$, ARMA(3,1)=$\mathcal{U}_2$

$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{1.06983e-05-7.8781e-06}{7.8781e-06} = 130.663$ reject ARMA(2,1)

(ii) Let ARMA(3,1)=$\mathcal{U}_1$, ARMA(4,1)=$\mathcal{U}_2$

$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{7.8781e-06-8.20009e-06}{8.20009e-06} = -14.3323$ choose ARMA(3,1)

3. Simulation

According to the Table 7.3.3, the ARMA(3,1) SPM concentration model is chosen as follows:

$$Y_n = A_1 Y_{n-1} + A_2 Y_{n-2} + A_3 Y_{n-3} + w_n + C_1 w_{n-1} \tag{7.38}$$

Set as:

$$\theta^T = [A_1, A_2, A_3, C_1] \tag{7.39}$$

$$x_n^T = [Y_{n-1}, Y_{n-2}, Y_{n-3}, e_{n-1}] \tag{7.40}$$

To solve equation (7.38), the RLSM (3.55)-(3.57) is used (substitute Y_n for y_n) in (3.55) and from (7.35), $p = 3, q = 1, m = 10, d = 40$. The time scale in here is $3\frac{3}{4}$ minutes per

run. Since the initial value are need, the first third data as our initial value we really start the model at n=4.

The computation procedure of ELSM is as follows:

(i) Construct x_n according to (7.34) and (7.36) ($n \geq 3$).
(ii) Select initial values of θ_3 and R_3.
(iii) Calculate K_n, R_n and θ_n according to (3.55)-(3.57) based on the K_{n-1}, R_{n-1} , θ_{n-1} and x_n ($n \geq 3$).

The simulation results are given as following:
(i) The five parameter matrices are:

$$A_1 = \begin{pmatrix} 0.914 & 0.291 & 0.222 & 0.007 & -0.041 & -0.108 & 0.026 & -0.095 & -0.043 & 0.089 \\ 0.308 & 0.655 & 0.222 & 0.124 & 0.011 & 0.027 & 0.002 & -0.066 & 0.006 & 0.024 \\ 0.168 & 0.175 & 0.556 & 0.167 & 0.135 & 0.112 & 0.048 & -0.012 & -0.004 & -0.041 \\ 0.024 & 0.109 & 0.165 & 0.551 & 0.199 & 0.165 & 0.121 & 0.033 & 0.006 & -0.094 \\ -0.008 & -0.003 & 0.137 & 0.190 & 0.567 & 0.209 & 0.154 & 0.104 & 0.050 & -0.071 \\ -0.103 & 0.033 & 0.099 & 0.180 & 0.220 & 0.560 & 0.236 & 0.112 & 0.039 & -0.070 \\ -0.048 & -0.043 & 0.008 & 0.122 & 0.185 & 0.215 & 0.530 & 0.165 & 0.148 & 0.043 \\ -0.044 & -0.013 & -0.022 & 0.052 & 0.150 & 0.103 & 0.205 & 0.532 & 0.158 & 0.169 \\ -0.036 & 0.057 & -0.033 & -0.016 & 0.043 & -0.011 & 0.188 & 0.131 & 0.613 & 0.324 \\ 0.053 & 0.032 & -0.051 & -0.114 & -0.069 & -0.100 & 0.127 & 0.135 & 0.311 & 0.947 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0.338 & -0.010 & 0.055 & -0.031 & -0.033 & -0.019 & 0.102 & -0.059 & -0.026 & 0.018 \\ 0.016 & 0.317 & 0.021 & 0.002 & -0.004 & -0.012 & 0.035 & -0.031 & -0.027 & 0.012 \\ 0.014 & -0.016 & 0.343 & 0.014 & -0.004 & 0.002 & 0.008 & -0.010 & 0.007 & 0.003 \\ 0.006 & -0.011 & 0.001 & 0.335 & 0.019 & -0.006 & 0.010 & -0.015 & 0.000 & 0.007 \\ 0.012 & -0.032 & 0.008 & 0.007 & 0.338 & 0.019 & -0.026 & -0.013 & 0.029 & -0.003 \\ 0.000 & 0.012 & -0.015 & 0.008 & 0.006 & 0.301 & 0.031 & -0.014 & 0.009 & -0.011 \\ 0.013 & -0.018 & -0.033 & 0.015 & 0.021 & 0.006 & 0.304 & -0.011 & 0.036 & -0.018 \\ -0.003 & 0.031 & -0.043 & -0.003 & 0.039 & 0.030 & 0.033 & 0.315 & 0.011 & -0.005 \\ -0.014 & 0.064 & -0.028 & -0.026 & 0.015 & -0.056 & 0.087 & -0.018 & 0.319 & -0.010 \\ -0.022 & 0.032 & 0.016 & -0.012 & 0.001 & -0.039 & 0.086 & -0.048 & -0.022 & 0.340 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} -0.254 & -0.332 & -0.120 & -0.074 & -0.018 & 0.076 & 0.191 & -0.013 & -0.007 & -0.049 \\ -0.301 & -0.042 & -0.189 & -0.125 & -0.015 & -0.044 & 0.079 & 0.019 & -0.044 & 0.020 \\ -0.156 & -0.219 & 0.120 & -0.144 & -0.148 & -0.116 & -0.039 & -0.017 & 0.013 & 0.043 \\ -0.010 & -0.133 & -0.165 & 0.111 & -0.166 & -0.185 & -0.108 & -0.069 & -0.010 & 0.109 \\ 0.042 & -0.056 & -0.118 & -0.176 & 0.108 & -0.175 & -0.211 & -0.136 & 0.001 & 0.057 \\ 0.099 & -0.010 & -0.133 & -0.164 & -0.208 & 0.043 & -0.170 & -0.136 & -0.014 & 0.059 \\ 0.093 & 0.016 & -0.068 & -0.085 & -0.140 & -0.202 & 0.080 & -0.185 & -0.070 & -0.076 \\ 0.042 & 0.075 & -0.070 & -0.063 & -0.078 & -0.170 & -0.145 & 0.092 & -0.141 & -0.183 \\ -0.007 & 0.062 & -0.030 & -0.041 & -0.015 & -0.100 & -0.009 & -0.161 & 0.038 & -0.331 \\ -0.105 & 0.029 & 0.076 & 0.083 & 0.062 & 0.015 & 0.042 & -0.229 & -0.340 & -0.241 \end{pmatrix}$$

$$C_1 = \begin{pmatrix} 0.523 & 0.118 & 0.101 & -0.013 & -0.026 & -0.091 & -0.005 & -0.013 & -0.015 & 0.054 \\ 0.118 & 0.400 & 0.133 & 0.074 & -0.033 & 0.015 & -0.029 & -0.049 & 0.015 & -0.001 \\ 0.041 & 0.107 & 0.314 & 0.094 & 0.065 & 0.073 & 0.022 & -0.023 & -0.011 & -0.025 \\ -0.013 & 0.064 & 0.088 & 0.310 & 0.107 & 0.114 & 0.050 & -0.011 & 0.003 & -0.069 \\ -0.012 & -0.018 & 0.092 & 0.096 & 0.315 & 0.101 & 0.102 & 0.044 & 0.008 & -0.069 \\ -0.082 & 0.014 & 0.073 & 0.062 & 0.124 & 0.312 & 0.132 & 0.041 & 0.012 & -0.057 \\ -0.020 & -0.028 & -0.015 & 0.046 & 0.094 & 0.132 & 0.316 & 0.086 & 0.036 & -0.006 \\ -0.005 & -0.041 & -0.000 & 0.018 & 0.071 & 0.058 & 0.115 & 0.308 & 0.044 & 0.080 \\ -0.016 & 0.046 & -0.023 & -0.001 & -0.000 & -0.017 & 0.066 & 0.068 & 0.339 & 0.192 \\ 0.076 & 0.034 & -0.057 & -0.089 & -0.044 & -0.053 & 0.025 & 0.119 & 0.145 & 0.566 \end{pmatrix}$$

ii)

TABLE 7.3.4. The mode of the parameter matrices in ARMA(3,1) SPM concentration model

	A_1	A_2	A_3	C_1
ARMA(3,1)	1.30509	0.431542	0.65197	0.768769

(iii)

MPEE=0.00883307 mg/l

MPVE=0.0209273 mg/l

MPV= 0.00255276

σ^2=7.8781 e-06

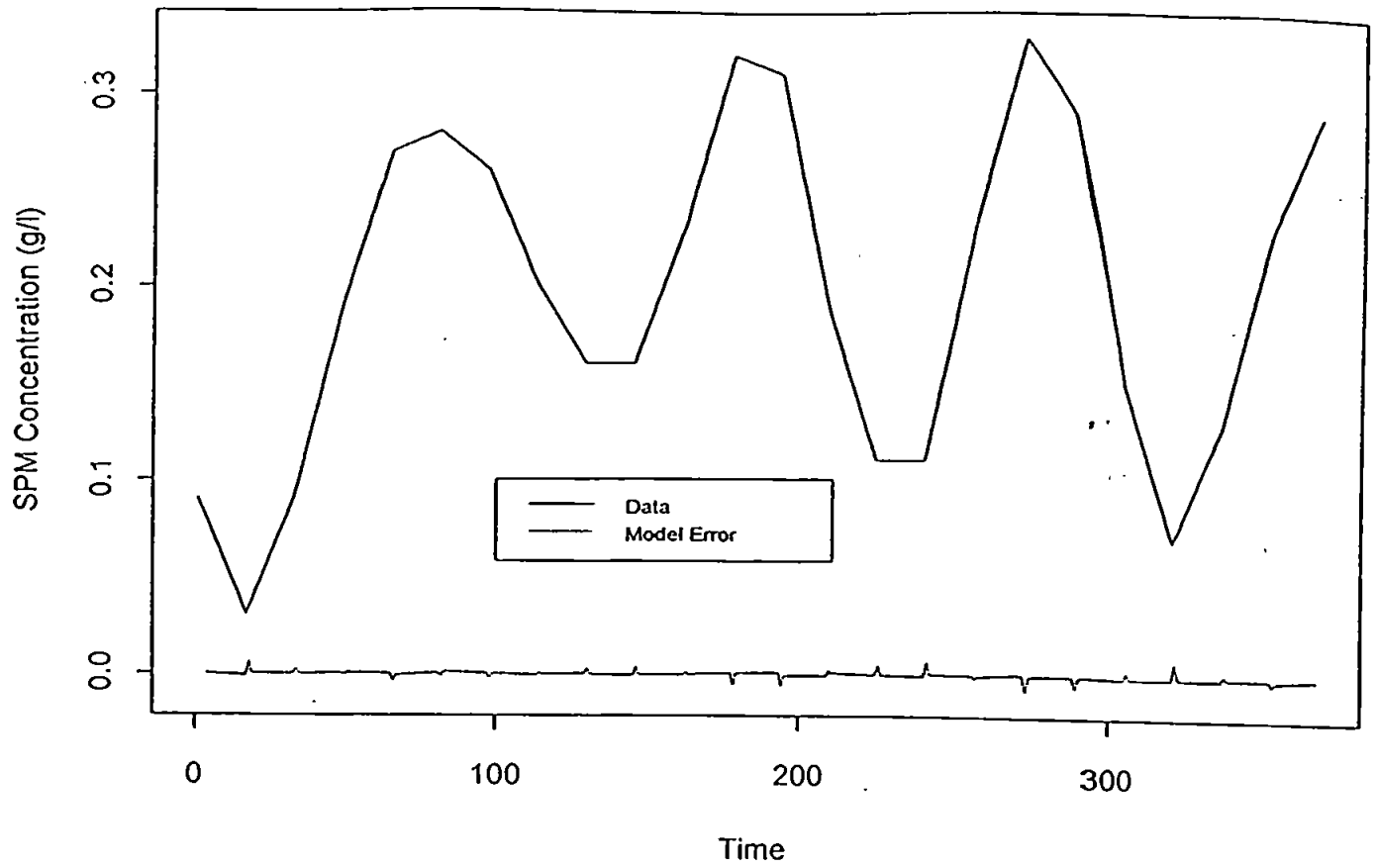


Figure 7.3.13. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.05H

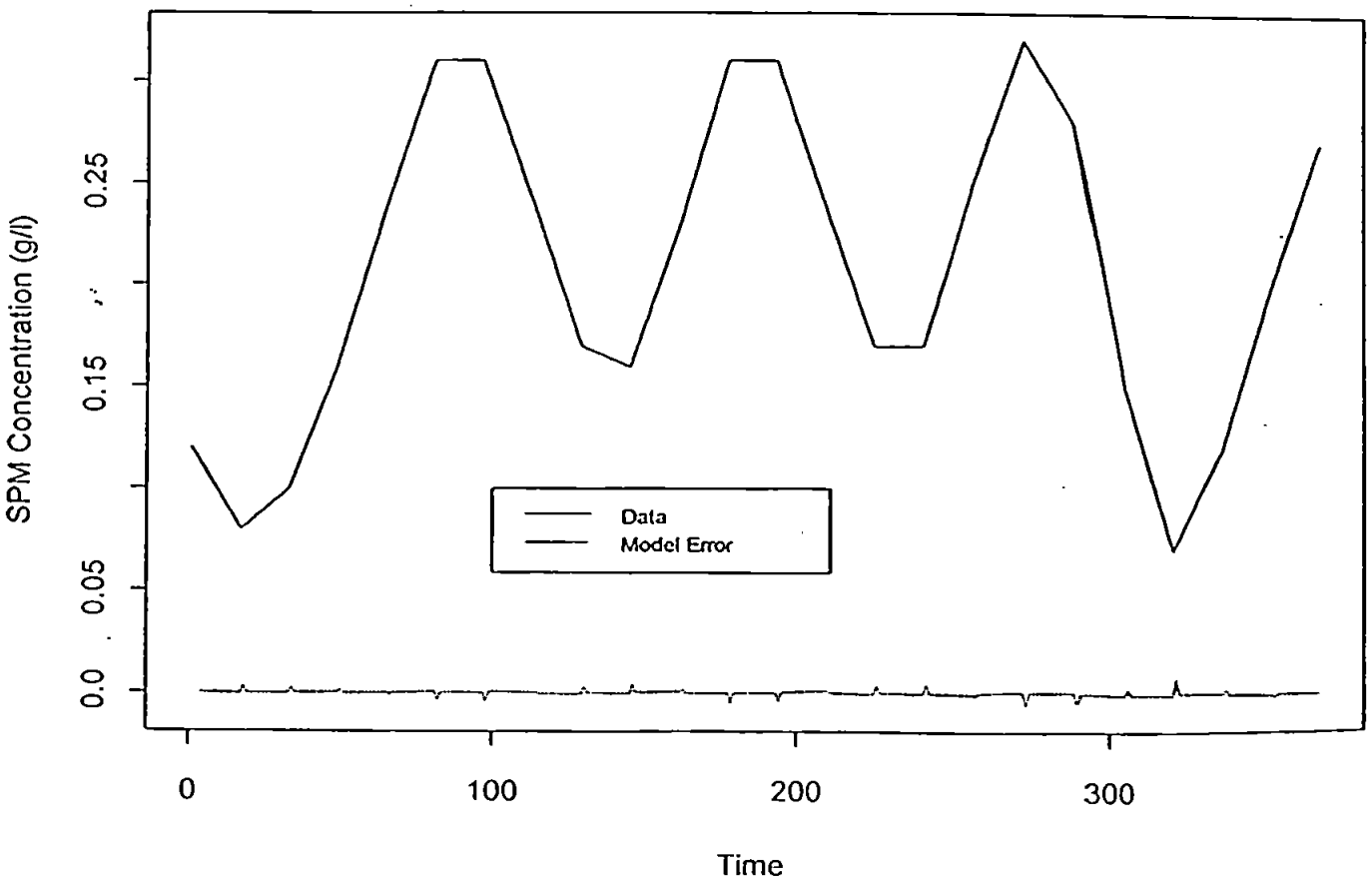


Figure 7.3.14. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.10H

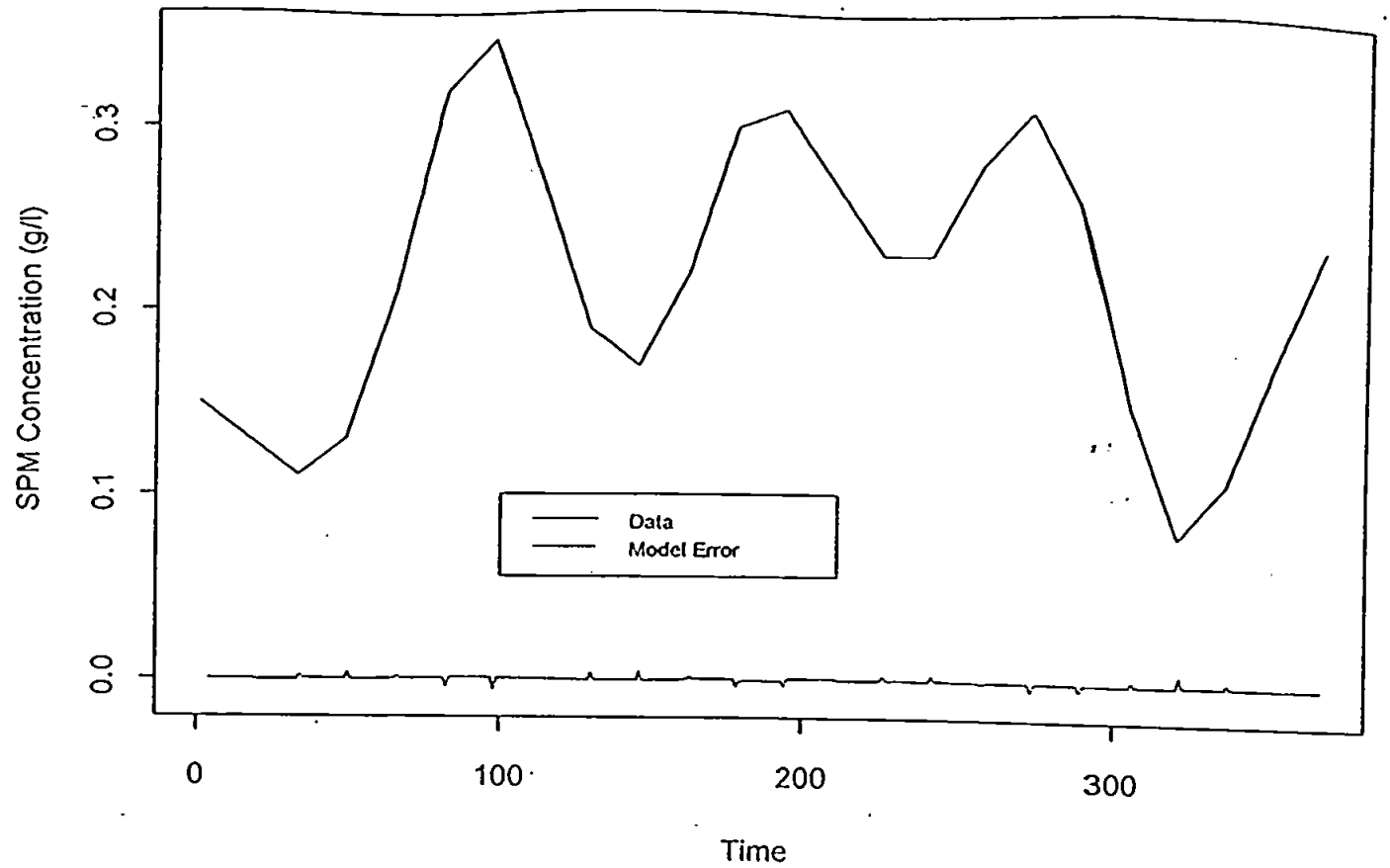


Figure 7.3.15. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.20H

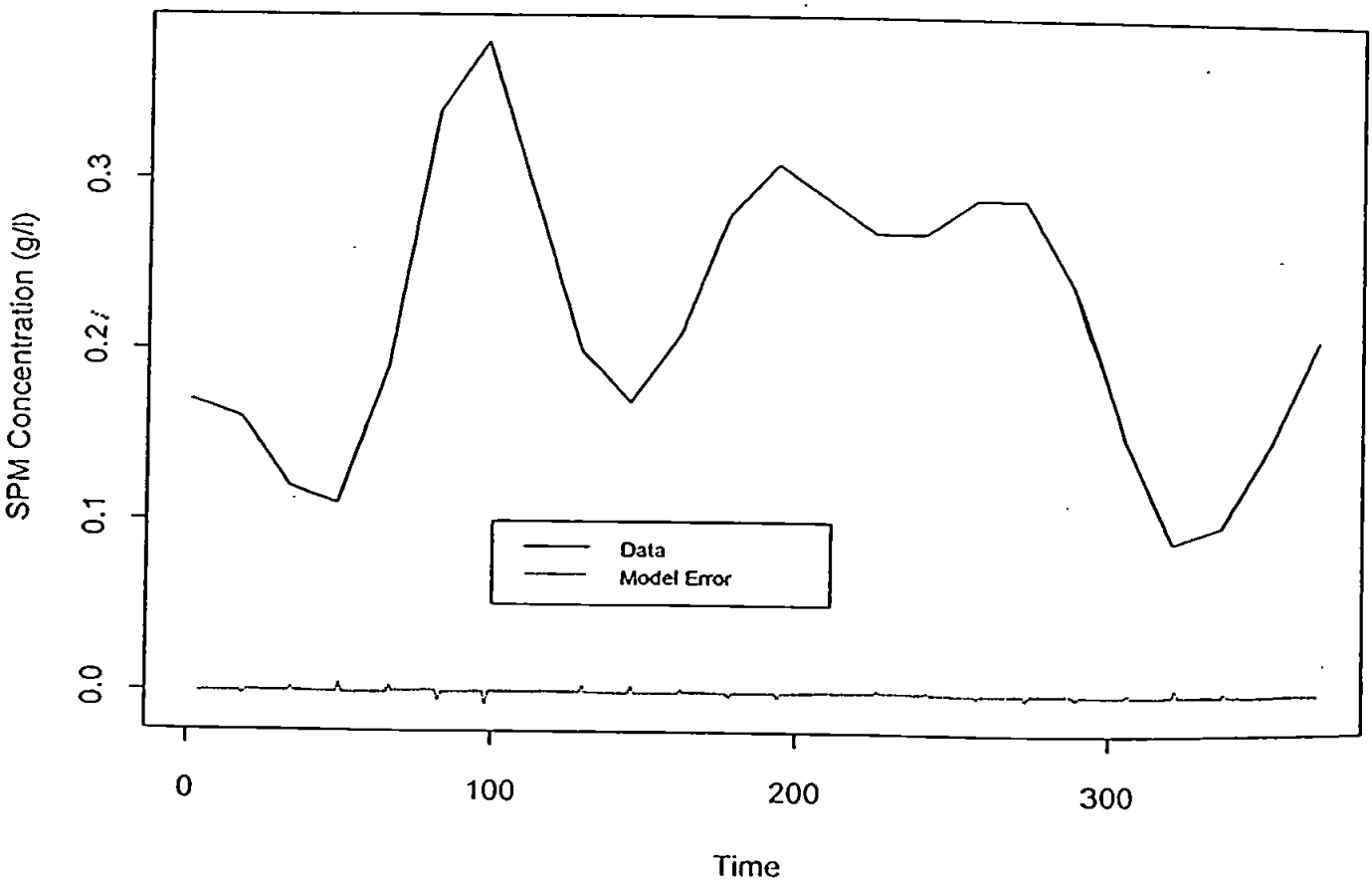


Figure 7.3.16. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.30H

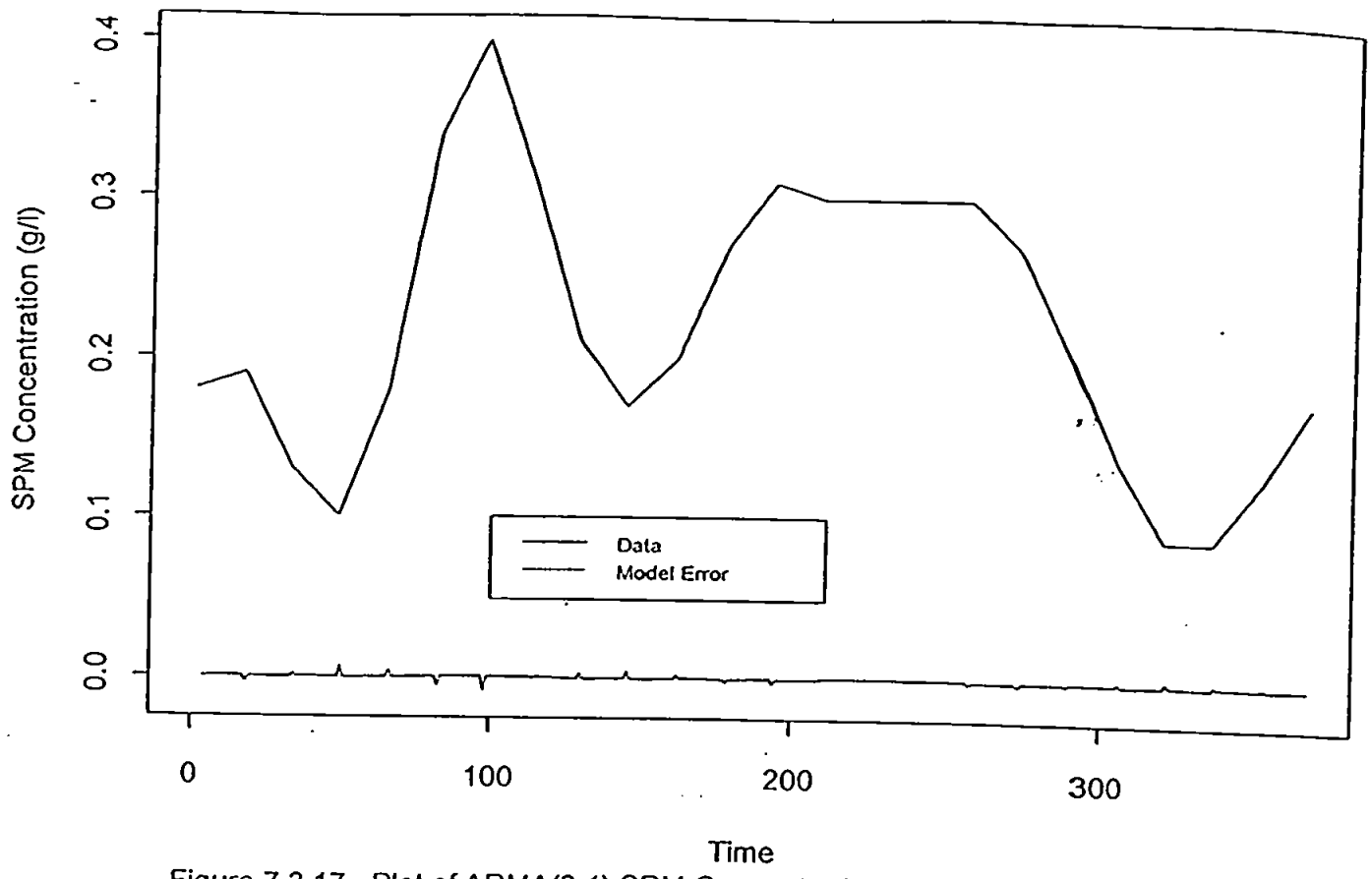


Figure 7.3.17. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.40H

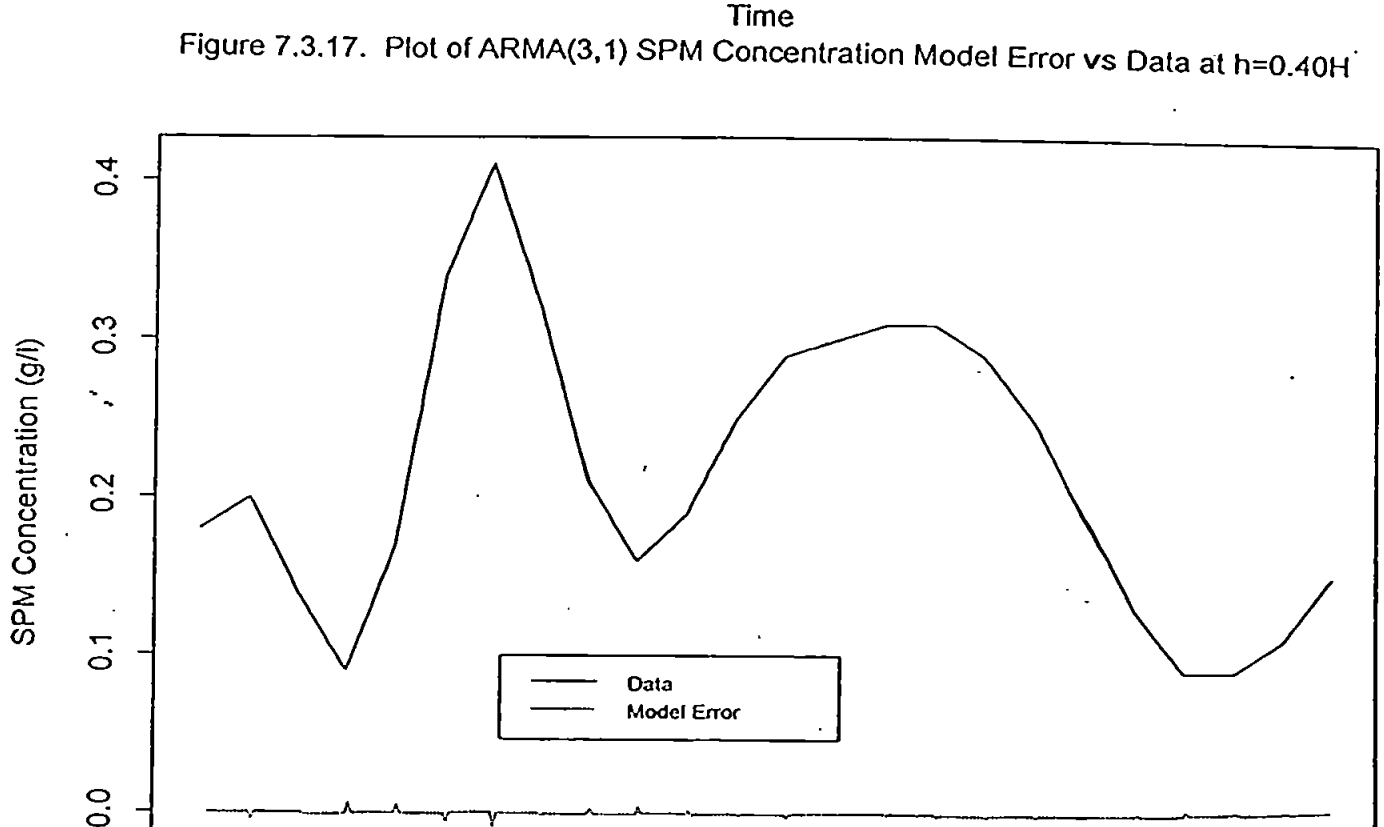


Figure 7.3.18. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.50H

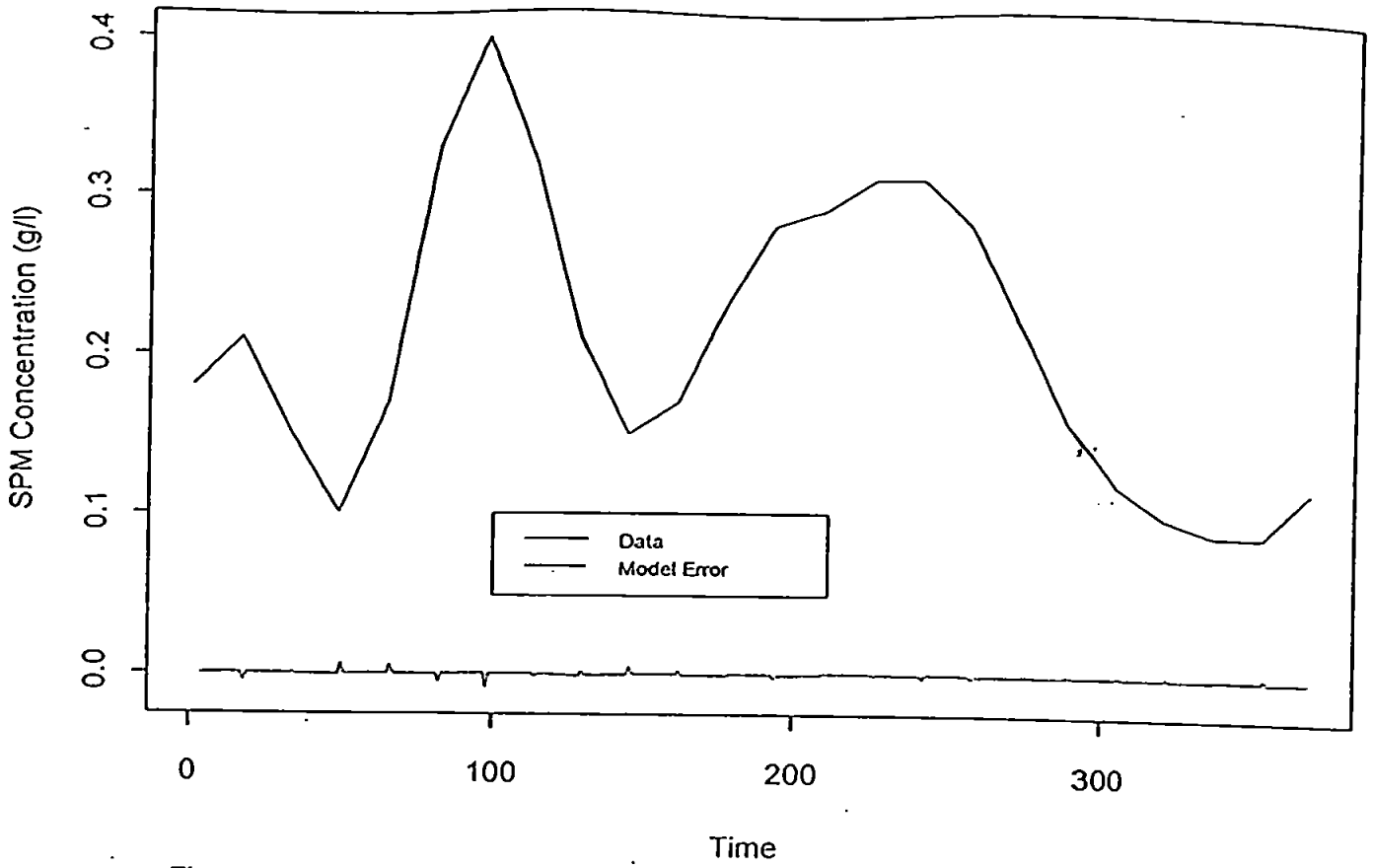


Figure 7.3.19. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.60H

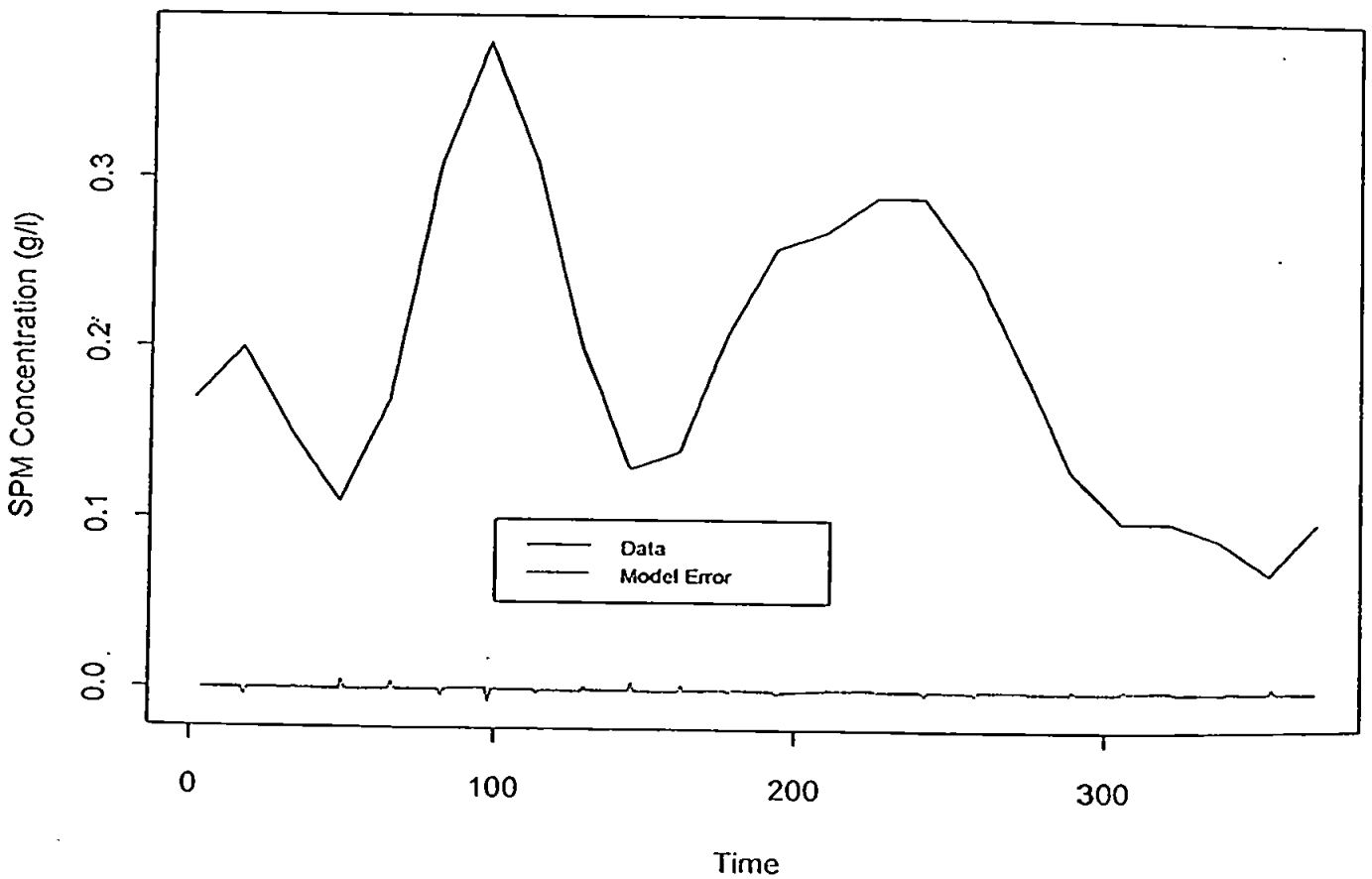


Figure 7.3.20. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.70H

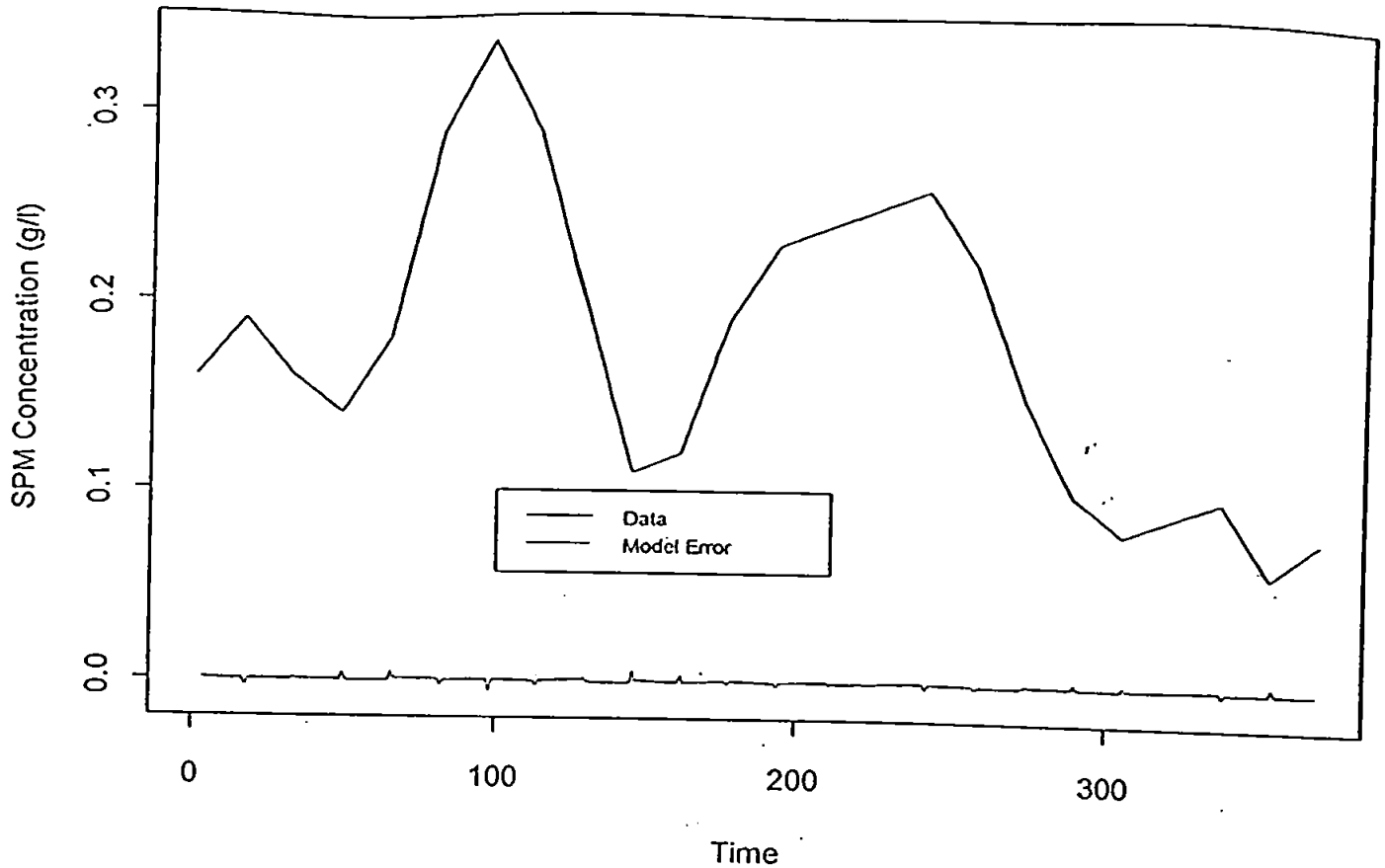


Figure 7.3.21. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.80H

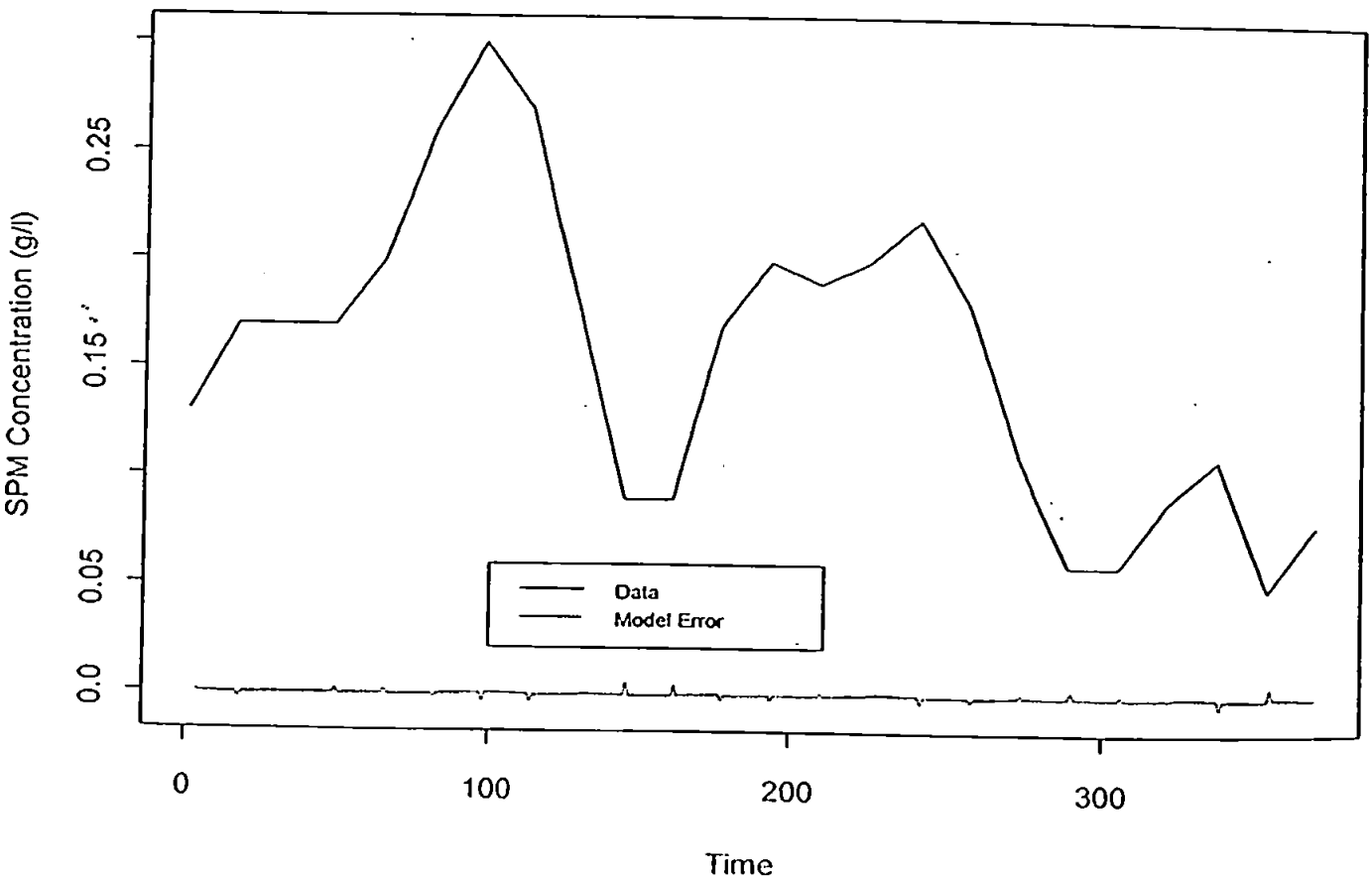


Figure 7.3.22. Plot of ARMA(3,1) SPM Concentration Model Error vs Data at h=0.90H

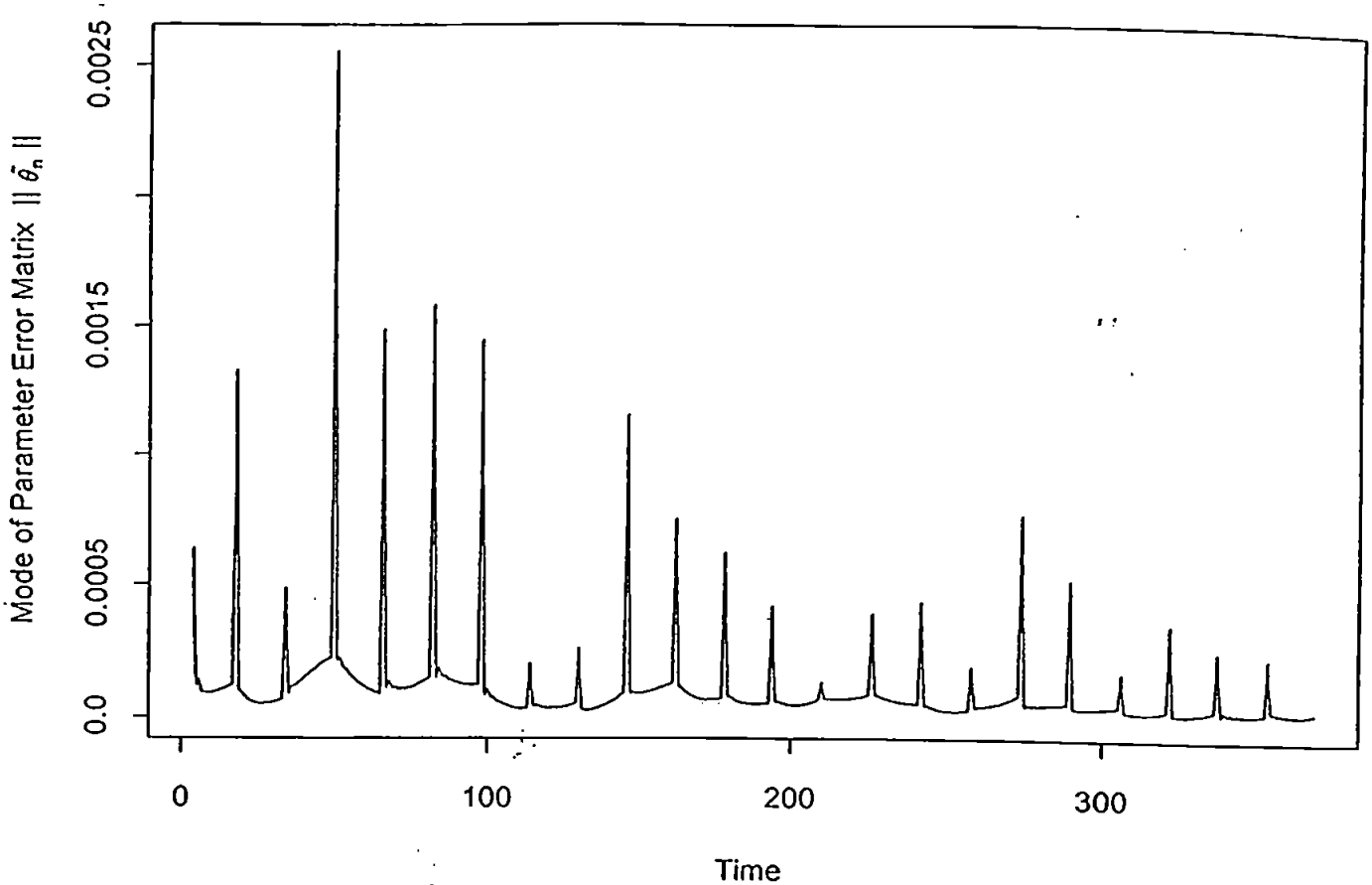


Figure 7.3.23. The variation of || $\tilde{\theta}_n$ || in ARMA(3,1) SPM Concentration Model with time

(iv) Figures 7.3.13–7.3.22 show the simulation of the suspended sediment concentration dynamics at different depth and Figure 7.3.23. gives the norm of the parameter matrix error dynamics in the SPM concentration model.
(v) The mode and the element of A_1, are comparatively large and the ones of A_2, A_3 are comparatively small which shows that the more recent the time, the more effect there is on the current variation.
(vi) Since A_1 is strongly diagonally dominant, we can say that the larger the distance between given layers, the less effect there is on the layer variation.

III. Model Comparison between AR(5) and ARMA(3,1)

The model comparison between AR(5) and ARMA(3,1) is given in Table 7.3.5.

TABLE 7.3.5. The Model Comparison

Model	MPEE	MPVE	MPV	σ^2
AR(5)	0.00947416	0.0221978	0.0249716	8.22566e-06
ARMA(3,1)	0.00883307	0.0209273	0.00255276	7.8781e-06

From Table 7.3.5, the MPEE, MPVE, MPV and σ^2 of the ARMA model are better than those in AR model. It is shown that the MPV of ARMA are much improved which means that the parameter matrices of ARMA time series model are 'nearly constant matrices'. Therefore the ARMA model presented here has less parameter identification matrices than the AR(5) model because it is a better description of the system, for data fitting and prediction then the AR model. This is because the ARMA model takes advantage of the information from the model error and estimation of system noise as well as under the assumption that the system noise is coloured noise.

7.3.2 ARMAX Model

In this subsection, according to the ocean science and geophysics, the suspended sediment concentration is closely related to the magnitude of the current velocity, so here the suspended sediment dynamics is taken as a unknown stochastic system. The current velocity profile is set as an input to the system model, and the suspended sediment concentration

is considered as the output of the system model. System Identification theory is applied in the model to identify the unknown parameter matrices of the model, based on in situ data collected in the study area. In this case, the system identification technique is applied to the Rufiji Delta, Tanzania, the data collected by (Fisher 1994) in March 11, 1993, to illustrate the quantitative relationship between suspended sediment concentration and current velocity. The aim is to show that this particular theory is suitable for matching known data sets, and this will be achieved by ensuring all the parameter matrices remain virtually constant when tackling future data sets. By operating the model in this way, the results obtained should simulate the data with some degree of accuracy.

The models in last subsection, used a suspended sediment concentration variation that was assumed to be dependent on is its own past through its temporal and spatial values as well as uncontrollable system noises. Here, the suspended sediment concentration profile is assumed to be related to the current velocity profile according to the physics and ocean geophysics. This appears more realistic and reasonable in an estuary environment. It is conceded therefore that the parameter matrices we find may depend on times and positions where the physical, chemical, microbiological and geographical processes remain not very well understood. The simulation results show the model developed has good agreement with the real data collected from the Rufiji Delta, Tanzania. The comparison of the multivariate model with univariate model is given to show that this type of model is a good one in as far as matching known data sets, and we shall show this by showing that all the elements of each matrix remain virtually constant when subjected to future data. This way, the model is shown to describe the data with accuracy and can be used for prediction.

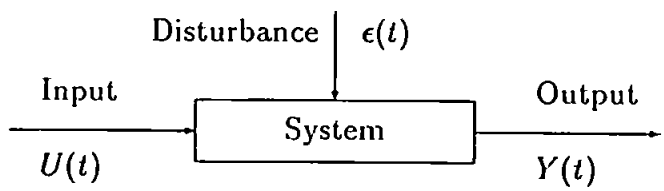


Fig 7.3.24. A Dynamic system with input $U(t)$, output $Y(t)$ and disturbance $\epsilon(t)$ where t denotes time.

We assume that the time series model here that describes the concentration of suspended sediment and current velocity is a discrete multivariable time-invariant stochastic linear system and can be represented by the following ARMAX(p,q,r) model:

$$A(z^{-1})Y_n + B(z^{-1})U_n = C(z^{-1})w_n \tag{7.41}$$

where

$$\begin{aligned} A(z^{-1}) &= I_m + A_1 z^{-1} + ... + A_p z^{-p} \\ B(z^{-1}) &= B_1 z^{-1} + ... + B_q z^{-q} \\ C(z^{-1}) &= I_m + C_1 z^{-1} + ... + C_r z^{-r} \end{aligned} \tag{7.42}$$

where w_n is the system noise and the restriction on it presented in (3.66) and (3.67), Y_n, U_n and w_n are m-dimensional vectors and z^{-1} is a unit delay operator and A_i, B_j and $C_k (i = 1, ..., p; j = 1, ...q; k = 1, 2, ..., r)$ are $m \times m$ unknown matrices to be estimated. I_m is an $m \times m$ unit matrix.
Set

$$\theta^T = [-A_1, ..., -A_p, B_1, ..., B_q, C_1, ..., C_r]_{m \times d} \tag{7.43}$$

$$x_n^T = [Y_{n-1}^T, ..., Y_{n-p}^T, U_{n-1}^T, ..., U_{n-q}^T, e_{n-1}^T, ..., e_{n-q}^T]_{1 \times d} \tag{7.44}$$

$$e_n = Y_n - \theta_n^T x_n \tag{7.45}$$

$$d = m \times (p + q + r) \tag{7.46}$$

here θ_n is the estimate of θ at time n and $[.]_{m \times d}$ and $[.]_{1 \times d}$ denote an $m \times d$ matrix and a $d-$dimensional row vector respectively.
It is easy to see that (7.41) also can be written by

$$Y_n = \theta^T x_n + C(z^{-1})w_n + e_n - C(z^{-1})e_n \tag{7.47}$$

We construct the vertical profile variable vector of suspended sediment concentration current velocity as the system output and it is a function of the water depth and time. Denote $Y_n(h)$ and $U_n(h)$ as the suspended sediment concentration and squared current velocity at height h and time n respectively (i.e. set water depth H, bottom $h = H$ and surface $h = 0$). So

$Y_n^T = [Y_n(h_1), Y_n(h_2), Y_n(h_3), ..., Y_n(h_m)]_{1\times m}$ and $U_n^T = [U_n(h_1), U_n(h_2), U_n(h_3), ..., U_n(h_m)]_{1\times m}$. where $0 \leq h_1 \leq h_2 \leq ... \leq h_m \leq H$.

In order to identify the system parameter matrix θ, we make use of the recursive algorithms (3.55–3.57).

2. Simulation

The simulations are consisted of two parts (i.e. order determination, and parameter identification). In here, we choose a multivariate time series model which is a simpler form of (7.41) as following ARMAX$(p, q, 1)$ model:

$$Y_n = A_1 Y_{n-1} + ... + A_p Y_{n-p} + B_1 U_{n-1} + ... + B_q U_{n-q} + w_n + C_1 w_{n-1}. \tag{7.48}$$

Set:

$$\theta^T = [A_1, ..., A_p, B_1, ..., B_q, C_1] \tag{7.49}$$

$$x_n^T = [Y_{n-1}, ..., Y_{n-p}, U_{n-1}, ..., U_{n-q}, e_{n-1}] \tag{7.50}$$

$$e_n = Y_n - \theta_n^T x_n. \tag{7.51}$$

2.1. Order determination

There is an identification of the model order problem here. The identification of a system order is a research branch in system identification which is very complicated. It is usually difficult to assess the order of the time series model. Some statistical methods and criteria are used to try to determine the order of the system, such as partial or inverse autocorrelation function and autocovariance function method, Akaike's final prediction error criterion, Akaike's information criterion and Parzen's autoregressive transfer function criterion etc. But unfortunately, all these criteria and methods may give more than one minimum, depend on assuming that the data are normally distributed, and sometimes indicate too many parameters (Chatfield 1980). Thus they should be used only as guides. So in this section we try to keep balance between the accuracy of the model and the number of parameters that need to be estimated.

The approach is to fit the model of progressively higher order, to calculate variance

of one-step prediction error σ^2 for each value of order (p, q, r), as well as to consider the MPVE, MPEE and MPV. The criterion is that if the addition of extra parameter matrices gives little improvement, we do not choose a higher order model.
The so-called F-test method (Söderström 1989) is used here to determine our model structure and the F-test results for our model candidates according to (4.11) and (4.12) are given as following in Table 7.3.6:

TABLE 7.3.6. The Order Comparison of the ARMAX Model

Model	MPEE	MPVE	MPV	σ^2
ARMAX(3,2,1)	0.00860554	0.0200166	0.0022704	9.58426e-06
ARMAX(4,2,1)	0.00822247	0.0195851	0.00238675	6.79282e-06
ARMAX(4,3,1)	0.00794057	0.0187704	0.002442882	6.42447e-06

Let ARMAX(3,2,1)=$\mathcal{U}_1$, ARMAX(4,2,1)=$\mathcal{U}_2$
$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{9.58462e-06-6.79282e-06}{6.79282e-06} = 149.993$
reject ARMAX(3,2,1)

Let ARMAX(4,2,1)=$\mathcal{U}_1$, ARMAX(4,3,1)=$\mathcal{U}_2$
$\chi^2_{0.05}(100) \approx 128.84$ and $x = 365 \times \frac{6.79282e-06-6.42447e-06}{6.42447e-06} = 20.9274$
choose ARMAX(4,2,1)

2.2. Parameter Estimation

From order determination we choose ARMAX(4, 2, 1) time series model for the Rufiji Delta as follows:

$$Y_n = A_1 Y_{n-1} + ... + A_4 Y_{n-4} + B_1 U_{n-1} + B_2 U_{n-2} + w_n + C_1 w_{n-1} \tag{7.52}$$

Set as:

$$\theta^T = [A_1, A_2, A_3, A_4, B_1, B_2, C_1] \tag{7.53}$$

$$x_n^T = [Y_{n-1}, ..., Y_{n-4}, U_{n-1}, U_{n-2}, U_{n-3}, e_{n-1}] \tag{7.54}$$

To solve equation (7.51), the RLSM (3.55)-(3.57) is used (substitute Y_n for y_n) in (3.55) and from (7.51), here $p = 4, q = 2, r = 1$ and $d = 70$ and the time scale here is 3.75

minutes per run. Since the initial values are needed, the first four data are taken as these initial values, so we really start the model at n=5.

The computation procedure of ELSM is as follows:

(i) Construct x_n according to (7.54) and (7.51) ($n \geq 4$).
(ii) Select initial values of θ_4 and R_4.
(iii) Calculate K_n, R_n and θ_n according to (3.55)-(3.57) based on the K_{n-1}, R_{n-1} , θ_{n-1} and x_n ($n \geq 4$).

The parameter estimation results are given as follows:
(i) The seven parameter matrices are:

$$A_1 = \begin{pmatrix} 0.926 & 0.203 & 0.139 & 0.048 & 0.036 & -0.008 & -0.008 & -0.110 & -0.076 & 0.058 \\ 0.228 & 0.654 & 0.225 & 0.176 & -0.007 & 0.131 & -0.030 & -0.144 & 0.005 & 0.051 \\ 0.128 & 0.203 & 0.493 & 0.199 & 0.152 & 0.177 & 0.014 & -0.066 & 0.008 & 0.010 \\ 0.013 & 0.104 & 0.146 & 0.590 & 0.198 & 0.183 & 0.145 & -0.041 & 0.049 & -0.035 \\ 0.035 & -0.055 & 0.167 & 0.264 & 0.389 & 0.207 & 0.320 & 0.056 & -0.015 & -0.014 \\ -0.090 & 0.052 & 0.111 & 0.191 & 0.156 & 0.535 & 0.307 & 0.107 & 0.006 & 0.001 \\ -0.013 & -0.066 & -0.024 & 0.125 & 0.231 & 0.254 & 0.516 & 0.203 & 0.109 & 0.031 \\ -0.017 & -0.066 & -0.004 & 0.068 & 0.093 & 0.133 & 0.269 & 0.559 & 0.076 & 0.187 \\ -0.059 & 0.038 & 0.021 & 0.090 & -0.009 & 0.007 & 0.198 & 0.084 & 0.630 & 0.292 \\ 0.030 & 0.012 & -0.019 & -0.035 & -0.029 & -0.029 & 0.053 & 0.184 & 0.251 & 0.857 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0.273 & -0.104 & -0.044 & -0.007 & -0.015 & 0.062 & 0.026 & -0.084 & -0.041 & 0.028 \\ -0.055 & 0.237 & -0.032 & 0.022 & -0.051 & 0.054 & -0.013 & -0.110 & -0.042 & 0.043 \\ -0.031 & -0.032 & 0.223 & 0.023 & -0.009 & 0.056 & -0.036 & -0.083 & -0.015 & 0.028 \\ -0.022 & -0.032 & -0.029 & 0.308 & -0.019 & 0.004 & 0.008 & -0.106 & -0.016 & 0.014 \\ 0.011 & -0.083 & 0.032 & 0.062 & 0.129 & -0.004 & 0.106 & -0.080 & -0.051 & 0.015 \\ -0.026 & -0.002 & -0.014 & 0.011 & -0.067 & 0.228 & 0.066 & -0.045 & -0.053 & 0.002 \\ 0.033 & -0.020 & -0.061 & -0.022 & -0.002 & 0.006 & 0.214 & -0.013 & -0.024 & -0.025 \\ 0.041 & 0.010 & -0.013 & 0.005 & -0.048 & -0.041 & 0.037 & 0.265 & -0.059 & -0.006 \\ 0.003 & 0.028 & -0.008 & 0.006 & -0.082 & -0.091 & 0.062 & -0.067 & 0.267 & 0.001 \\ 0.019 & 0.001 & -0.012 & -0.019 & -0.038 & -0.065 & -0.034 & -0.022 & -0.040 & 0.311 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} -0.093 & -0.216 & -0.105 & -0.002 & -0.039 & 0.128 & 0.045 & -0.065 & 0.005 & 0.045 \\ -0.152 & -0.022 & -0.170 & -0.042 & -0.030 & 0.030 & 0.036 & -0.051 & -0.067 & 0.054 \\ -0.079 & -0.163 & 0.043 & -0.075 & -0.106 & -0.016 & -0.055 & -0.091 & -0.033 & 0.019 \\ -0.029 & -0.114 & -0.128 & 0.116 & -0.140 & -0.087 & -0.057 & -0.127 & -0.069 & 0.032 \\ -0.023 & -0.085 & -0.044 & -0.059 & -0.040 & -0.124 & -0.029 & -0.165 & -0.063 & 0.019 \\ 0.034 & -0.015 & -0.066 & -0.069 & -0.179 & 0.033 & -0.073 & -0.122 & -0.078 & -0.016 \\ 0.022 & 0.015 & -0.054 & -0.084 & -0.121 & -0.120 & 0.038 & -0.118 & -0.074 & -0.042 \\ -0.001 & 0.032 & -0.037 & -0.025 & -0.126 & -0.134 & -0.093 & 0.080 & -0.089 & -0.100 \\ 0.026 & 0.012 & -0.022 & -0.035 & -0.092 & -0.110 & 0.027 & -0.099 & 0.045 & -0.134 \\ 0.003 & 0.003 & 0.016 & 0.027 & 0.002 & -0.038 & -0.027 & -0.091 & -0.151 & 0.006 \end{pmatrix}$$

$$A_4 = \begin{pmatrix} -0.037 & -0.053 & -0.005 & 0.064 & -0.060 & 0.143 & -0.010 & -0.113 & 0.007 & 0.060 \\ 0.073 & -0.025 & -0.135 & 0.003 & 0.049 & 0.038 & 0.083 & 0.001 & -0.091 & 0.075 \\ 0.072 & -0.094 & 0.050 & -0.002 & -0.047 & 0.047 & 0.035 & -0.012 & 0.024 & 0.044 \\ 0.089 & -0.043 & -0.057 & 0.107 & -0.078 & -0.001 & 0.037 & -0.011 & -0.016 & 0.118 \\ 0.018 & 0.033 & 0.044 & 0.011 & -0.003 & -0.043 & 0.023 & -0.089 & 0.051 & 0.096 \\ 0.044 & 0.008 & -0.019 & 0.007 & -0.104 & 0.048 & -0.004 & -0.012 & 0.046 & 0.068 \\ -0.023 & 0.079 & 0.051 & 0.005 & -0.053 & -0.048 & 0.066 & -0.042 & 0.020 & 0.030 \\ -0.023 & 0.079 & 0.051 & 0.005 & -0.053 & -0.060 & -0.045 & 0.069 & 0.027 & -0.059 \\ 0.001 & 0.002 & 0.005 & 0.006 & -0.000 & -0.008 & 0.126 & 0.008 & -0.028 & -0.136 \\ -0.076 & -0.026 & 0.035 & 0.087 & 0.078 & 0.042 & 0.054 & -0.058 & -0.137 & -0.131 \end{pmatrix}$$

$$B_1 = \begin{pmatrix} -0.030 & 0.020 & -0.014 & 0.053 & -0.098 & 0.032 & -0.133 & -0.038 & -0.050 & -0.002 \\ -0.011 & -0.025 & 0.006 & 0.065 & -0.041 & 0.071 & -0.213 & 0.103 & -0.017 & -0.070 \\ -0.012 & -0.029 & -0.004 & 0.015 & -0.019 & 0.049 & -0.167 & 0.060 & -0.006 & -0.008 \\ 0.017 & -0.051 & -0.024 & -0.002 & -0.053 & -0.003 & -0.088 & 0.015 & 0.029 & 0.006 \\ -0.010 & -0.003 & -0.060 & 0.036 & 0.008 & -0.026 & -0.077 & 0.088 & 0.002 & 0.001 \\ 0.052 & -0.040 & -0.075 & 0.023 & 0.027 & -0.025 & -0.102 & 0.128 & -0.002 & -0.020 \\ 0.012 & -0.005 & -0.120 & 0.147 & 0.034 & -0.022 & -0.055 & 0.075 & -0.024 & 0.019 \\ 0.012 & 0.027 & -0.099 & 0.058 & 0.042 & -0.081 & -0.054 & 0.074 & -0.016 & 0.001 \\ 0.004 & 0.026 & -0.113 & 0.082 & 0.027 & -0.021 & -0.030 & 0.058 & 0.017 & 0.000 \\ 0.022 & 0.012 & -0.116 & 0.060 & 0.016 & -0.017 & -0.069 & 0.092 & 0.032 & -0.027 \end{pmatrix}$$

$$B_2 = \begin{pmatrix} -0.052 & 0.058 & 0.026 & 0.114 & -0.047 & 0.077 & -0.046 & 0.062 & 0.063 & -0.004 \\ 0.023 & -0.040 & 0.007 & 0.076 & 0.017 & 0.046 & -0.189 & 0.107 & 0.032 & 0.018 \\ 0.009 & -0.009 & 0.033 & 0.056 & 0.030 & 0.058 & -0.144 & 0.081 & 0.021 & -0.038 \\ -0.002 & 0.003 & 0.030 & 0.070 & 0.056 & 0.027 & -0.056 & 0.011 & 0.060 & -0.087 \\ 0.030 & 0.003 & -0.036 & 0.054 & 0.002 & -0.008 & -0.094 & 0.063 & 0.009 & -0.027 \\ 0.002 & -0.023 & -0.016 & 0.090 & 0.038 & -0.021 & -0.140 & 0.097 & -0.020 & 0.003 \\ 0.023 & 0.005 & -0.081 & 0.136 & -0.043 & -0.072 & -0.078 & 0.076 & -0.020 & -0.020 \\ 0.017 & 0.016 & -0.101 & 0.067 & 0.048 & -0.061 & -0.040 & 0.100 & -0.048 & 0.034 \\ 0.036 & 0.021 & -0.124 & 0.063 & 0.004 & -0.049 & -0.036 & 0.025 & 0.022 & -0.051 \\ 0.024 & 0.040 & -0.130 & 0.066 & 0.036 & 0.026 & -0.120 & 0.073 & -0.025 & -0.019 \end{pmatrix}$$

$$C_1 = \begin{pmatrix} 0.526 & 0.087 & 0.058 & -0.060 & 0.023 & -0.071 & -0.006 & -0.060 & -0.010 & 0.041 \\ 0.087 & 0.361 & 0.144 & 0.042 & -0.056 & 0.034 & -0.061 & -0.127 & 0.060 & 0.030 \\ 0.028 & 0.116 & 0.272 & 0.070 & 0.065 & 0.050 & -0.021 & -0.055 & 0.011 & -0.006 \\ -0.023 & 0.071 & 0.073 & 0.264 & 0.072 & 0.057 & 0.049 & -0.032 & 0.028 & -0.007 \\ 0.018 & -0.041 & 0.045 & 0.125 & 0.183 & 0.085 & 0.156 & 0.028 & -0.030 & -0.017 \\ -0.079 & 0.008 & 0.054 & 0.053 & 0.061 & 0.276 & 0.157 & 0.001 & 0.005 & -0.009 \\ -0.001 & -0.058 & -0.065 & 0.059 & 0.103 & 0.103 & 0.276 & 0.070 & 0.052 & -0.039 \\ 0.026 & -0.048 & -0.013 & 0.003 & 0.028 & 0.051 & 0.148 & 0.299 & -0.001 & 0.073 \\ -0.026 & -0.003 & -0.013 & 0.059 & -0.033 & -0.019 & 0.035 & 0.014 & 0.364 & 0.123 \\ 0.040 & -0.023 & -0.044 & -0.020 & -0.066 & -0.055 & -0.025 & 0.088 & 0.131 & 0.485 \end{pmatrix}$$

(ii)

MPEE=0.00822247 mg/l

MPVE=0.0195851 mg/l

MPV=0.00238675

σ^2=6.79282e-06

(iii) Figures 7.3.25–7.3.34 show the simulation of the suspended sediment concentration dynamics at different depth and Figure 7.3.35. gives the norm of the parameter matrix error dynamics in the suspended sediment concentration model.

(iv) Since A_1 and A_2 are strongly diagonally dominant, it is shown that the larger the distance between given layers, the less effect is there on the layer variation for suspended sediment concentration.

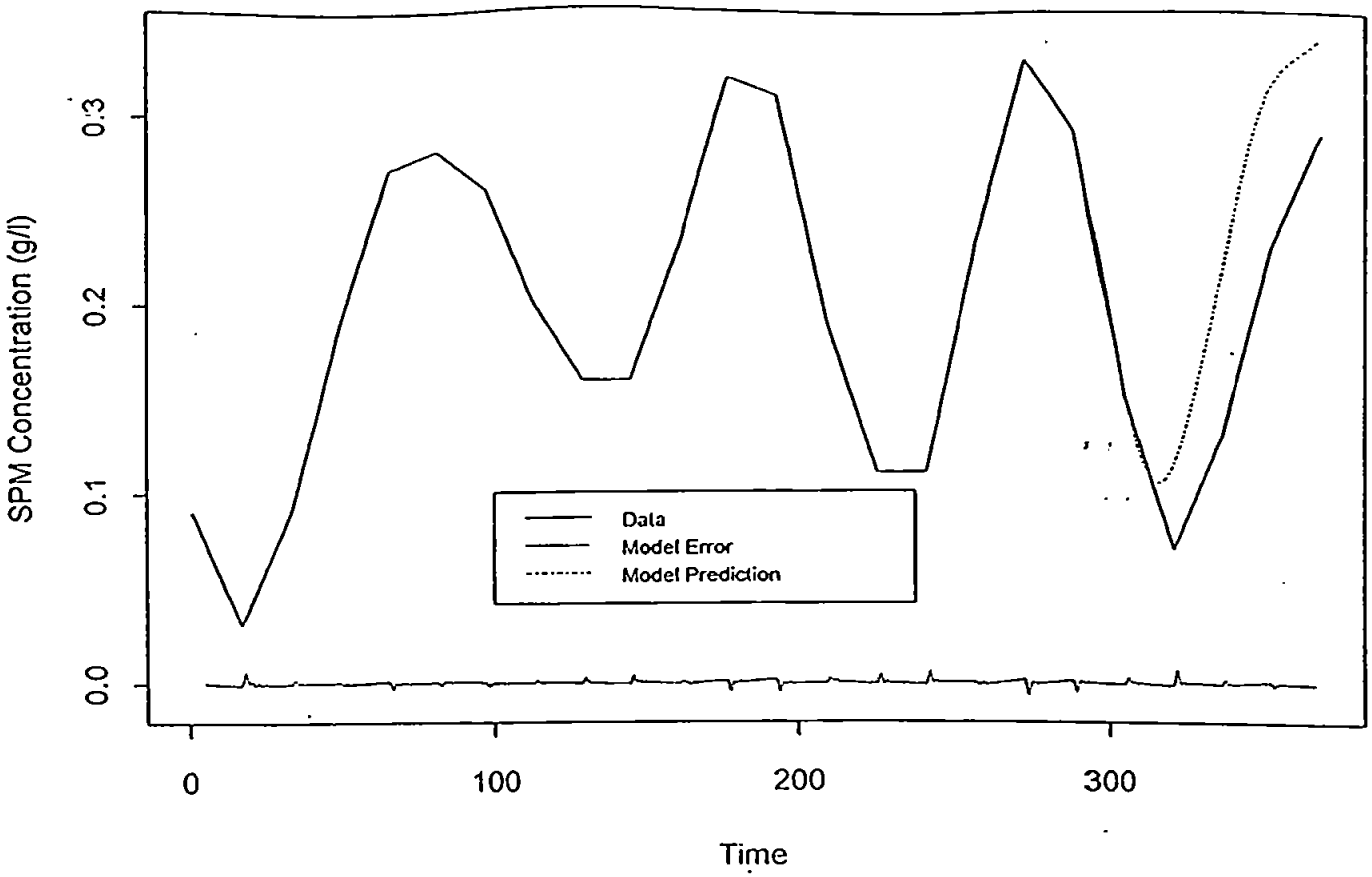


Figure 7.3.25. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.05H

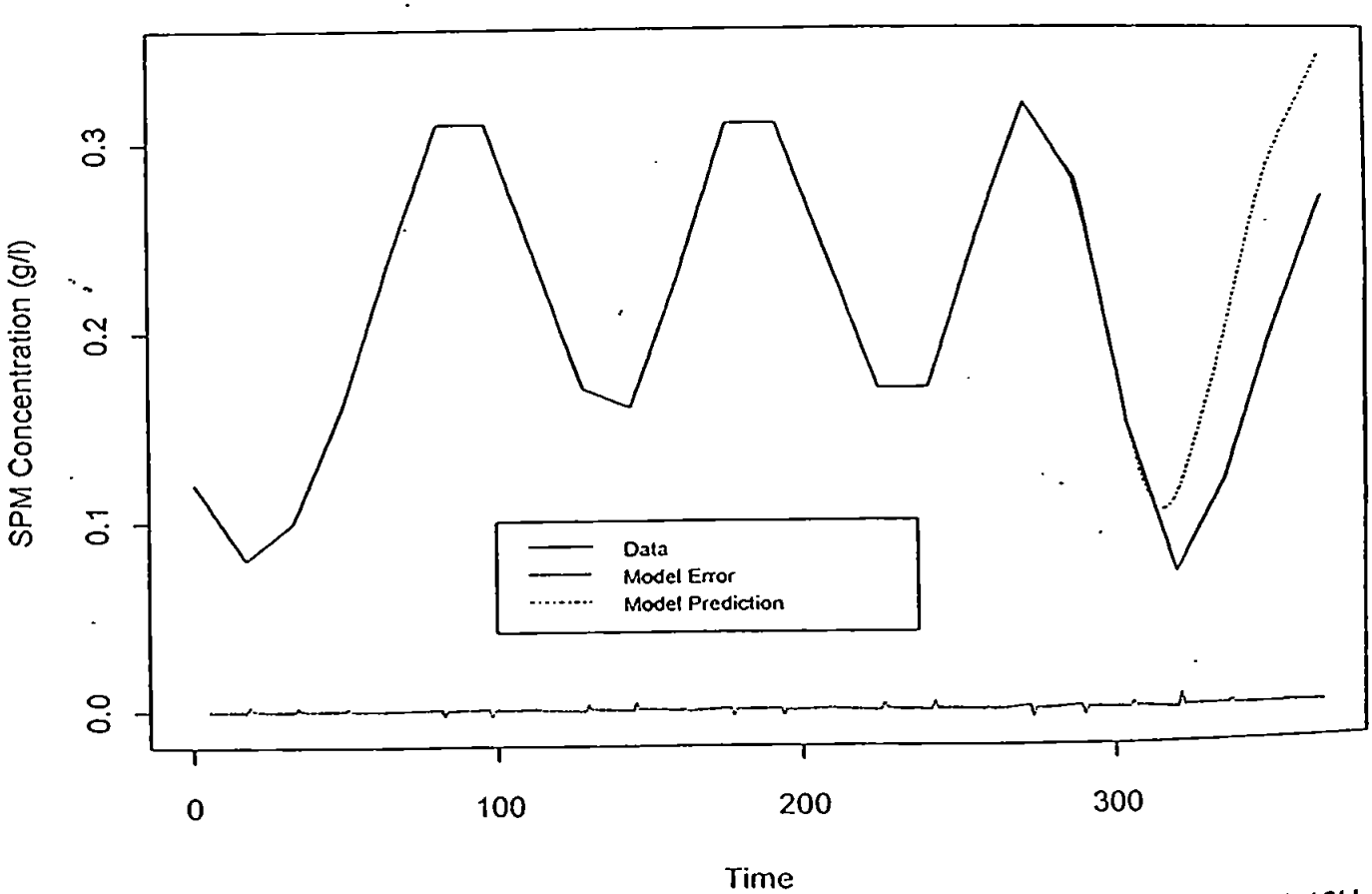


Figure 7.3.26. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.10H

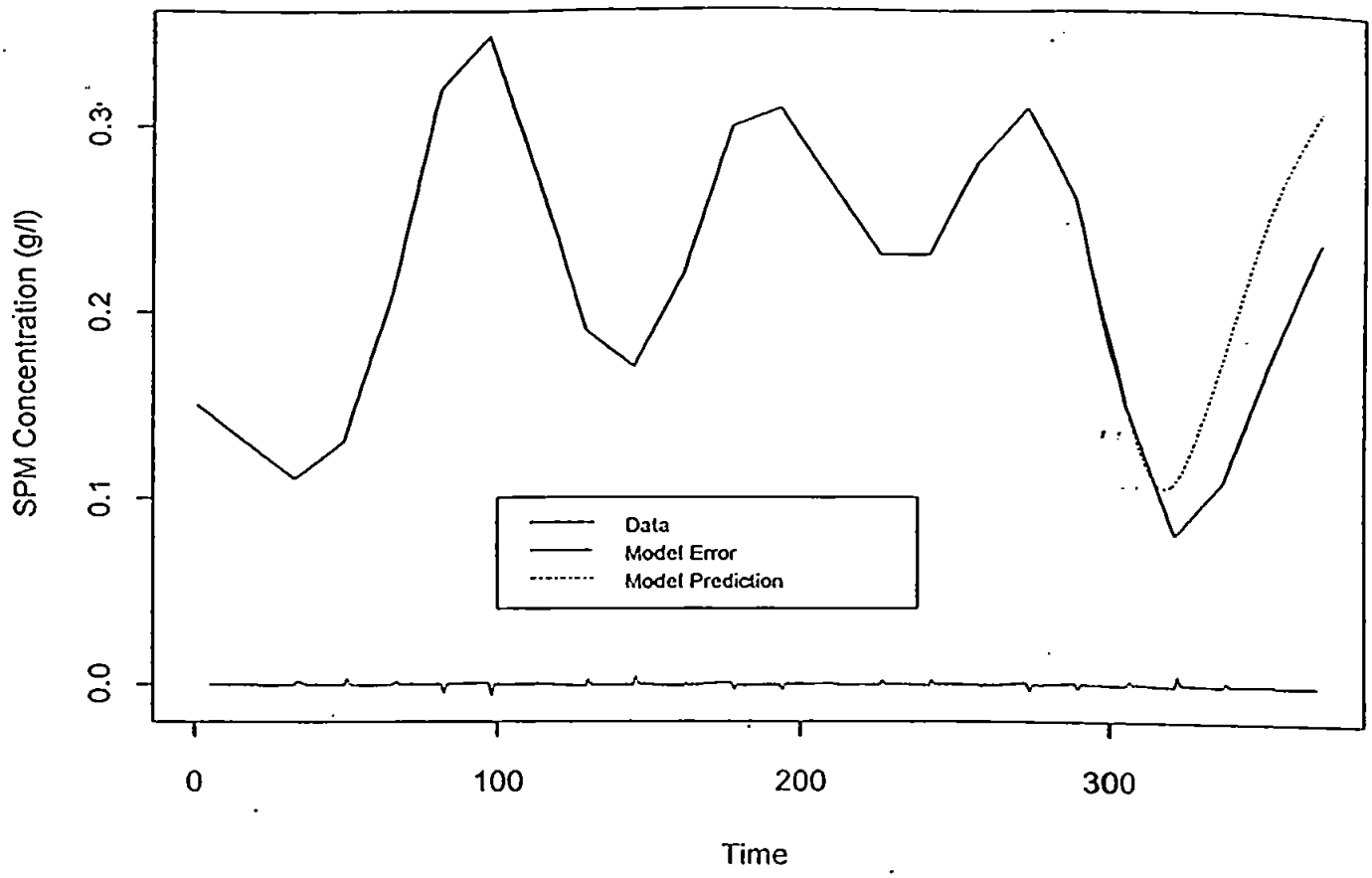


Figure 7.3.27. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.20H

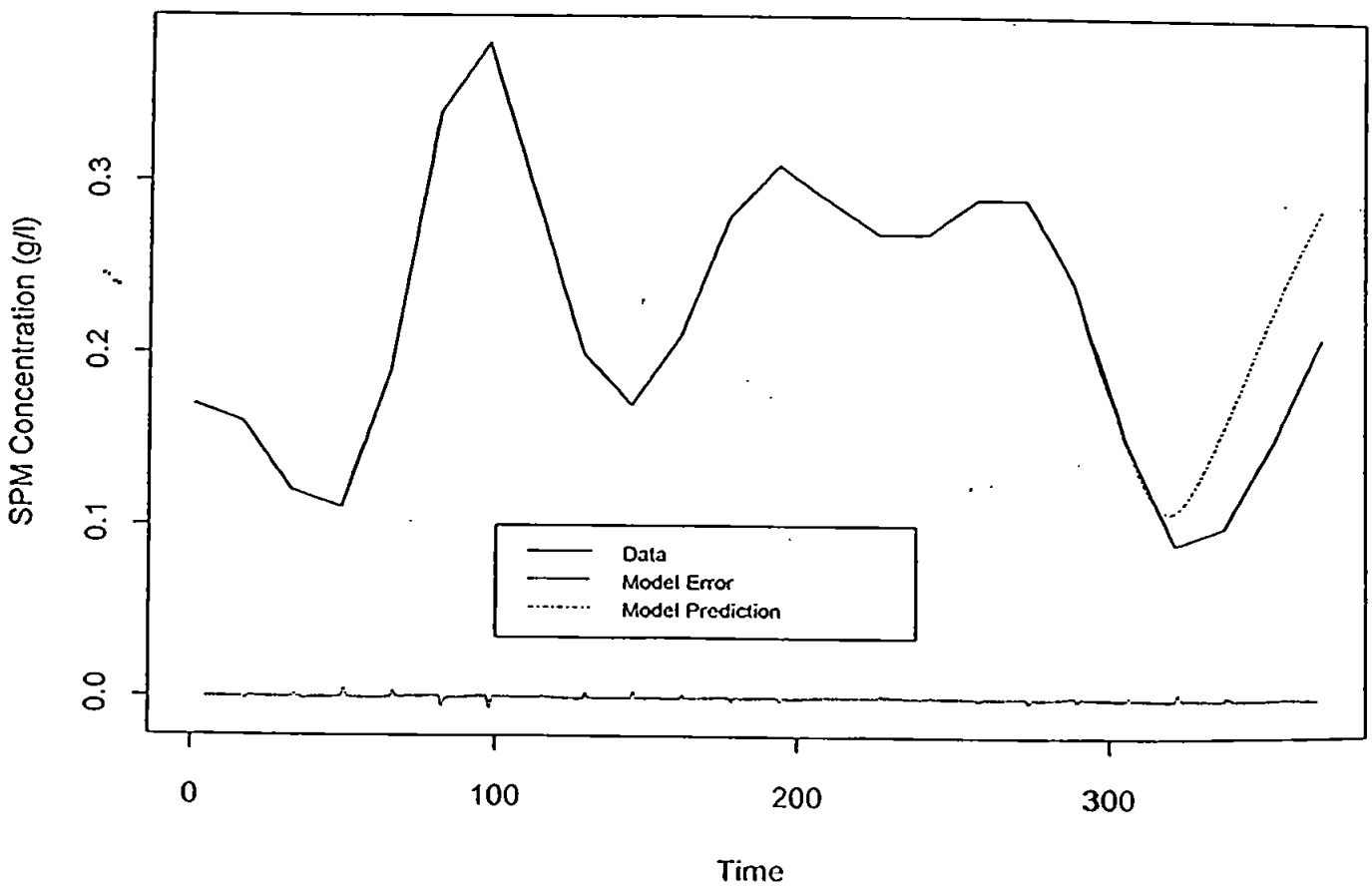


Figure 7.3.28. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.30H

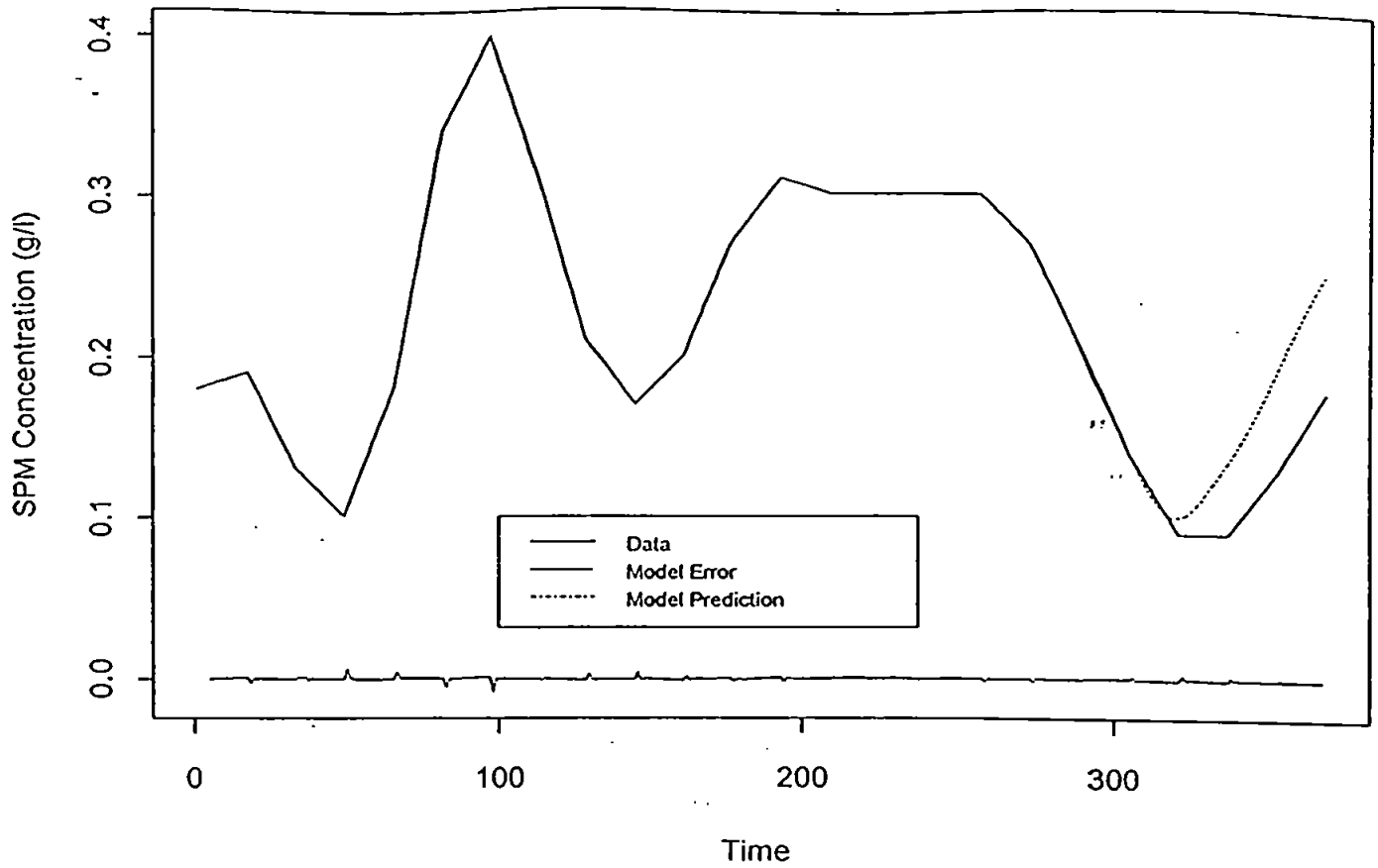


Figure 7.3.29. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.40H

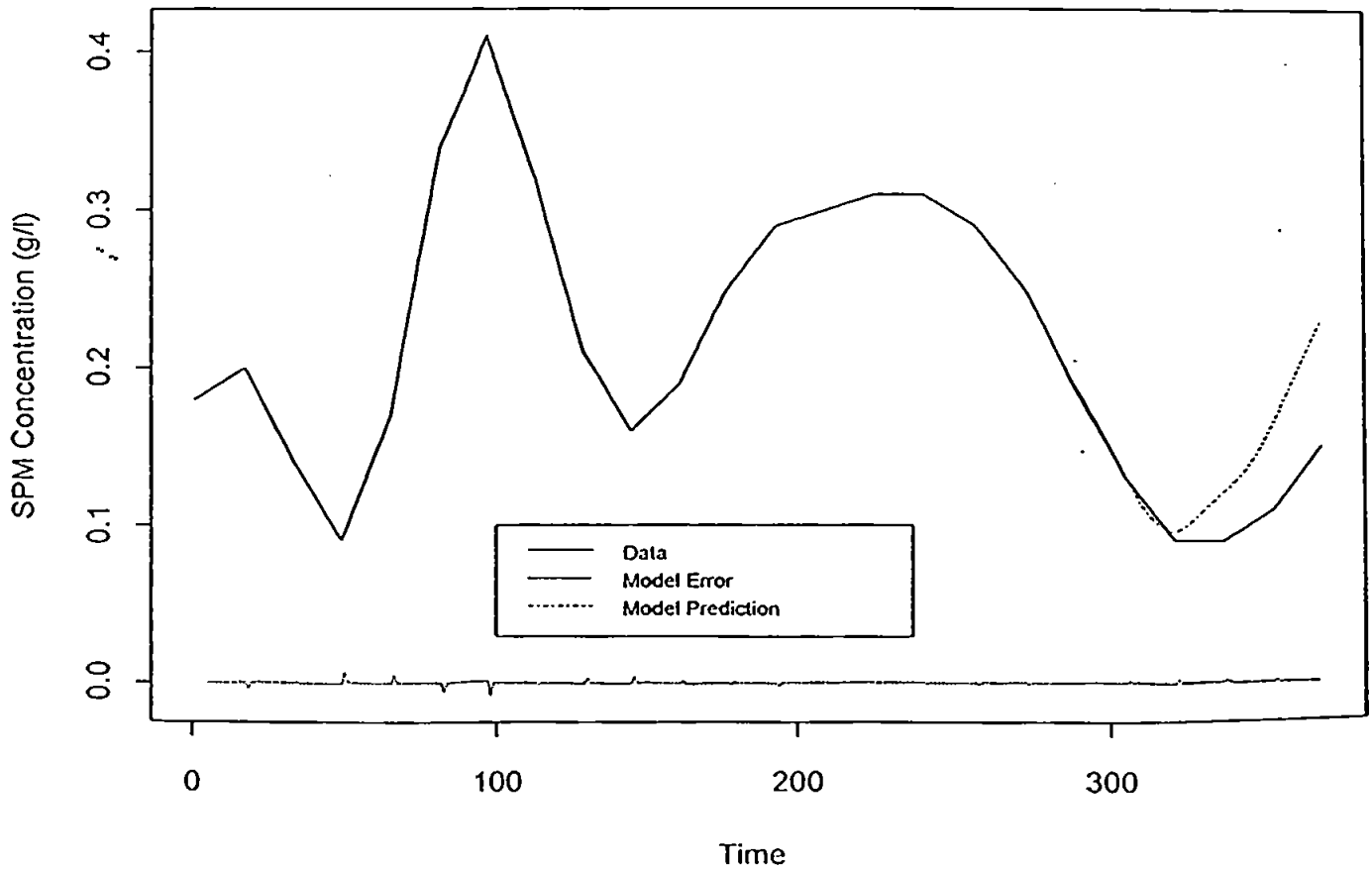


Figure 7.3.30. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.50H

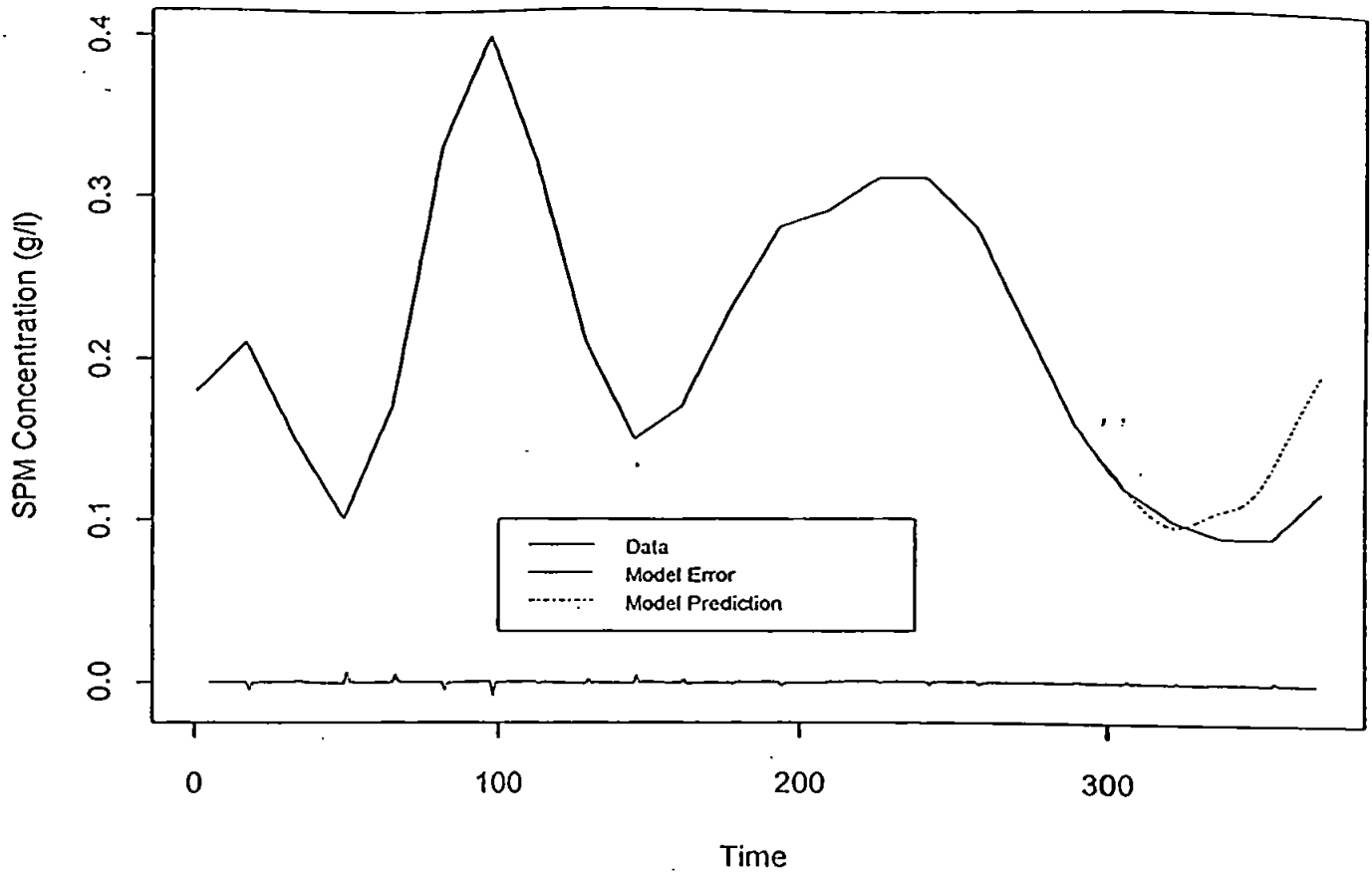


Figure 7.3.31. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.60H

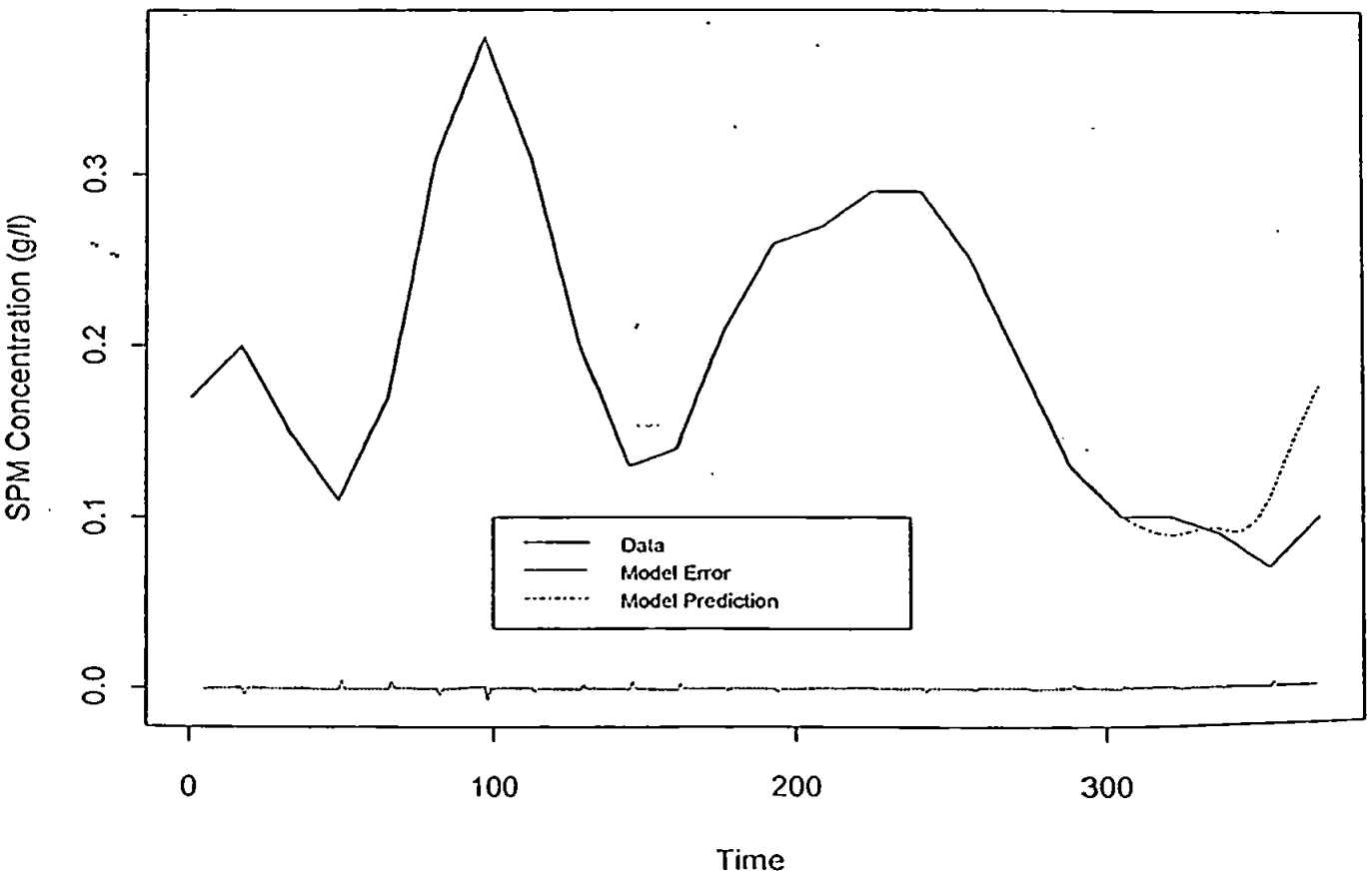


Figure 7.3.32. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.70H

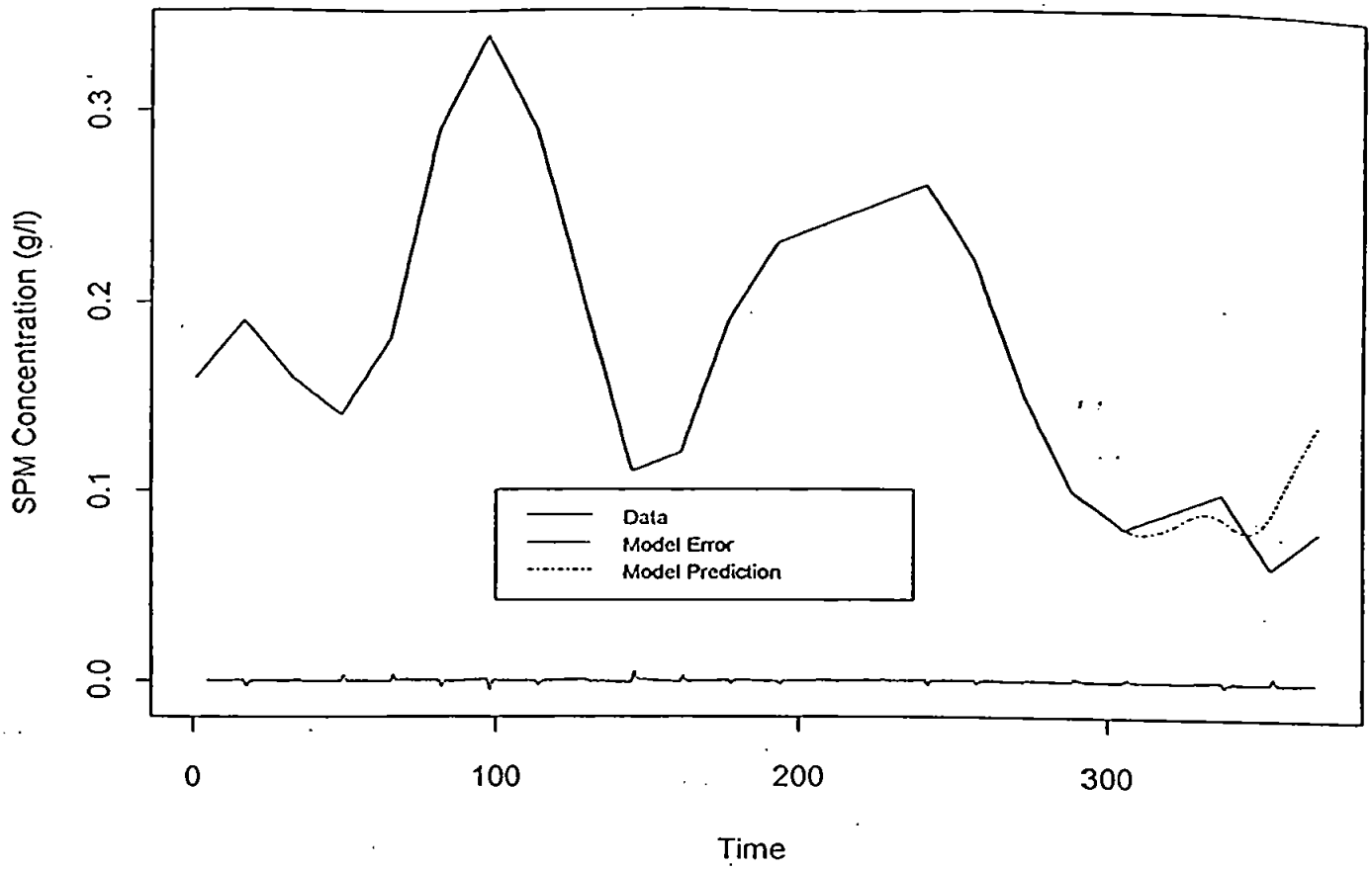


Figure 7.3.33. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.80H

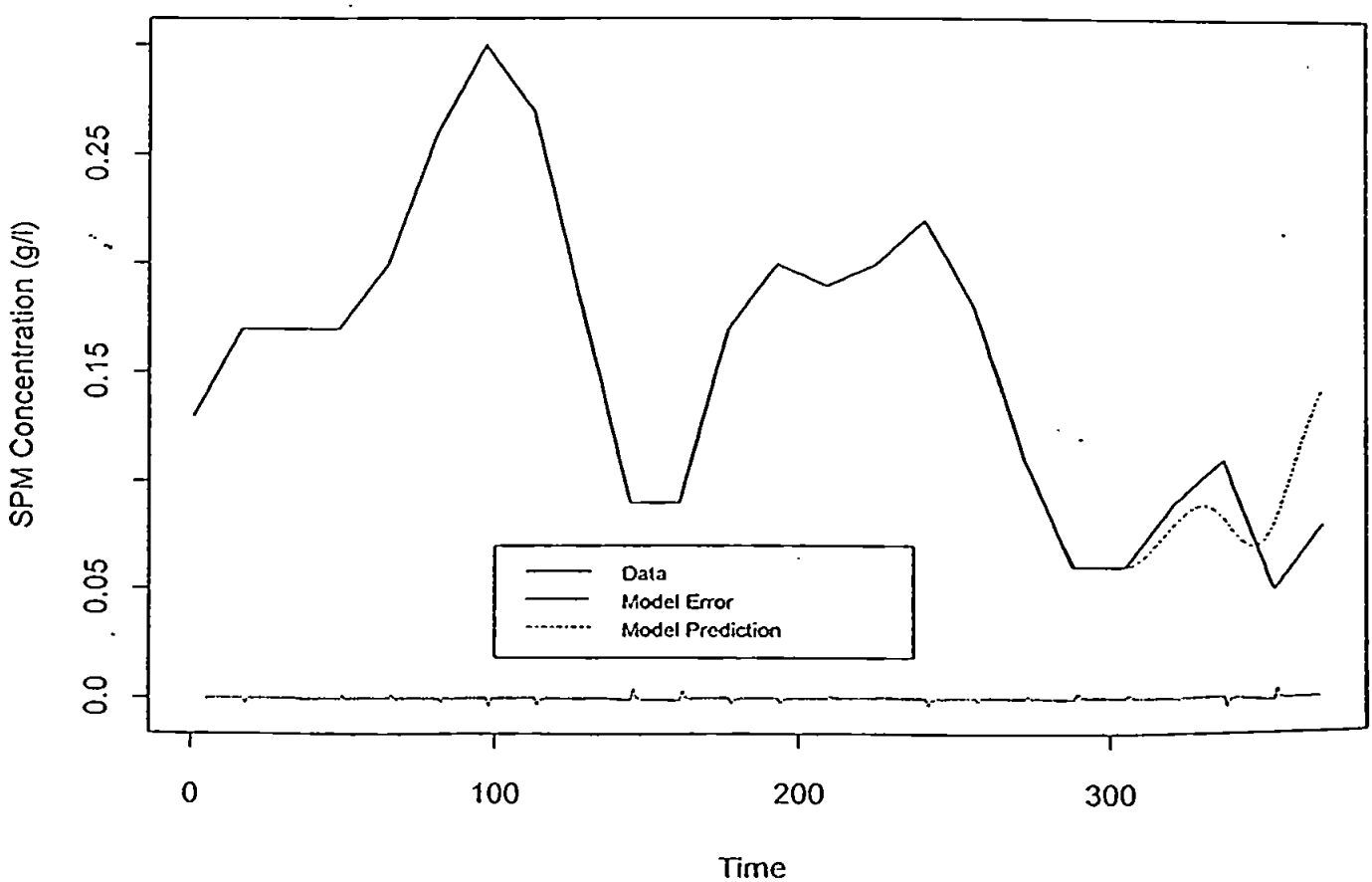


Figure 7.3.34. Plot of ARMAX(4,2,1) Model Error, Model Prediction vs Data at h=0.90H

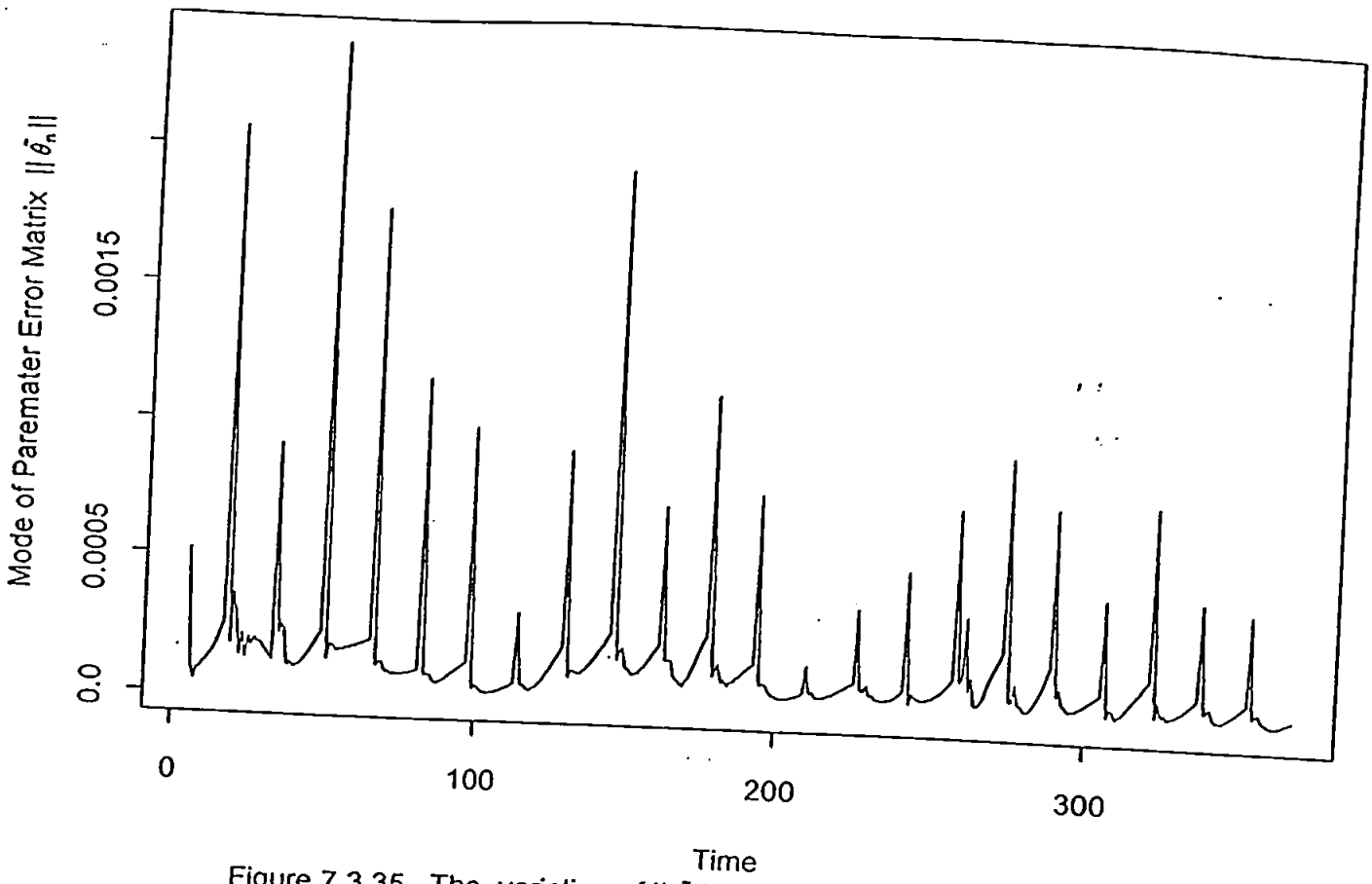


Figure 7.3.35. The variation of $\|\tilde{\theta}_n\|$ in ARMAX(4,2,1) Model with the time

3. Long term prediction of ARMAX (4,2,1)

The time series model we presented here not only can be used in data fitting and short time prediction as shown in last subsection but also can be used in comparatively long term prediction of the SPM sediment dynamics. If the time series model arises and its parameter matrices are known, it is easy to use to form predictions. Define $f_{n,1}$ as the optimum one-step prediction of Y_{n+1}. Here, Y_{n+1} will be generated by ARMAX (4,2,1) model as follows:

$$Y_{n+1} = A_1 Y_n + ... + A_4 Y_{n-3} + B_1 U_n + B_2 U_{n-1} + w_{n+1} + C_1 w_n.$$

The noise term w_{n+1} is unknown, so taking the expectation of both sides of the above mentioned equation and w_n is estimated by $f_{n-1,1} - Y_n$ (if the Y_n is unknown, $f_{n-1,1}$ is the estimate of Y_n, so the estimate of w_n is zero.) the optimum one-step prediction $f_{n,1}$ is

$$f_{n,1} = A_1 Y_n + ... + A_4 Y_{n-3} + B_1 U_n + B_2 U_{n-1} + C_1(f_{n-1,1} - Y_n)$$

similarly, Y_{n+2} will be generated by

$$Y_{n+2} = A_1 Y_{n+1} + (A_2 Y_n + ... + A_4 Y_{n-2} + B_1 U_{n+1} + B_2 U_n + C_1(f_{n,1} - Y_{n+1})) + w_{n+2}$$

The term w_{n+2} is unknown, where as the term in parentheses is entirely known at time n, and the first term is predicted by $A_1 f_{n,1}$, so the optimum two-step prediction $f_{n,2}$

$$f_{n,2} = A_1 f_{n,1} + (A_2 Y_n + ... + A_4 Y_{n-2} + B_1 U_{n+1} + B_2 U_n).$$

It is obvious how further predictions are formed: One simply writes down the generating mechanism for the value to be predicted, with everything that is known part of the prediction and everything that is not known replaced by its optimum prediction so that we can get the optimum k-step prediction $f_{n,k}, k = 1, 2, ...$ based on the values $Y_n, Y_{n-1}, ...$ The ARMAX (4,2,1) model in (7.52) are used here for long term prediction. We start at run n=307, under the assumption that we do not know the value of $Y_n, n \geq 307$ and get the series $f_{307,1}, f_{307,2}, ..., f_{307,64}$ up to 64-step prediction (4 hour ahead prediction). The simulation results and comparison with the data set are shown in Figures 7.3.25–7.3.35. The simulation results show that the maximum estimate error is

$$\max_{1 \leq k \leq 64} \| f_{307,k} - Y_{n+307} \| \leq 0.09.$$

The predictions are within the reasonable range of the real data set which shows the very good long term prediction property of the ARMAX model. From Figures 7.3.25–7.3.35, the prediction can describe trends and variations of the different layer of SPM dynamics

quite well although there is a little bit of an underestimate of the down shots at about n=318 in the layers from h=0.05H to h=0.4H which start to make a large discrepancy with the data. The simulations here also shown that the shorter the time, the better the prediction in common with all other models.

7.3.3 Comparison with the Univariate Model

We have presented three different kinds of time series SPM concentration model in the last subsection. Two of them (AR(5) and ARMA(3,1)) are the univariate model and the multivariate model ARMAX (4,2,1). The model coimparison between AR(5), ARMA(3,1) and ARMAX (4,2,1) is given in Table 7.3.7.

TABLE 7.3.7 The Model Comparison

Model	MPEE	MPVE	MPV	σ^2
AR(5)	0.00947416	0.0221978	0.0249716	8.22566e-06
ARMA(3,1)	0.00883307	0.0209273	0.00255276	7.8781e-06
ARMAX(4,2,1)	0.00822247	0.0195851	0.00238675	6.79282e-06

From Table 7.3.7, the MPEE, MPVE, MPV and σ^2 of the ARMAX model are better than those in AR and ARMA model. Therefore the multivariate model presented here is the best description of the system, including data fitting and prediction. It is confirmed that the SPM concentration profile dynamics do have dynamic relationships with the current velocity dynamics in the Rufiji Delta, Tanzania and the reason that ARMAX model works better than other models is due to the fact that it takes advantage of the information from the current velocity profiles.

7.4 Conclusion and Discussion

What we have presented here is an alternative to traditional current velocity and suspended sediment dynamical models. The equations that describe the behaviour of the

sediment dynamics have been replaced by a time series model containing some parameter matrices. The reality of the model is justified by using data to determine the parameters through a recursive time series procedure. To those familiar with traditional hydraulic modelling, it may seem unusual that we use seven parameter matrices with many free parameters, but all of them turn out to be virtually constant. This shows that the model fits the data very well, and hence we contend that it contains a good representation of the physics within, in this case, Rufiji Delta, Tanzania from which the data were taken. It is therefore a type of inverse modelling method. The models obtained by system identification here have the following properties, in contrast to traditional mathematical modelling:

(a) The model parameters have limited validity (They are valid for a certain working point, a certain type of coast, certain season etc.), but the model structure seems good since it fits the data very well.
(b) They give little physical, chemical and biological insight since in most cases the parameters of the model have no direct physical, chemical and biological meaning for the time being and the parameters are used here only as tools to give a good description of the dynamic system's overall behaviour. Further research is in progress to find relationships between our parameters and measurable quantities.
(c) They are relatively easy to construct and use.
(d) They have greater short term accuracy, especially for complex situation, than traditional hydraulic models.

Chapter 8

Conclusion

Two kinds of the modelling of sediment transport processes have been studied. The system identification theory provides a background for the use of development of a satisfactory description of current velocity and suspended sediment concentration dynamics. The strong consistency and convergence rate of recursive least squares method for the univariate and multivariate one variable and multivariable model, especially for the coloured noise case derived from chapter 3, gives the theoretical guarantee for the model description and accuracy.

The distinguishing character of our time series model is that it can be easily changed to an on-line or real-time identification method. *i.e.* If no new data collected, the model predictions can be used to describe the current velocity and SPM concentration dynamics and if there are new data collected, the new information is taken into the model for verifying and modifying the model parameter or parameter matrices and lets the model be adaptive to real dynamics in time. Since the natural world is always changing, storms, typhoons and other natural phenomenon are quite unpredictable in the long term. You can not expect an unchangeable model to work very well for a natural changing world in the long term. So it is very important to take the latest information, updated data to revise and renew the model for adapting the real current velocity and SPM concentration dynamics. The time series modelling technique presented in this thesis provides a novel and practical method to modelling sediment transport dynamics.

For current velocity modelling, from the results given in the chapter 6 and chapter 7,

the AR or ARMA model are suitable for description, data fitting and forecasting both in the estuary and nearshore regions. The models present a good periodical and trend character of the current velocity dynamics.

For the suspended sediment concentration model presented in the chapter 6 and chapter 7, more than one time series are considered in our model and the quantitative relationship are set up. The strong consistency and convergence rate of recursive least squares method given in chapter 3 provide the theoretical background for the model description of the system and guarantee the modelling work. Since the current velocity profile data are economic and easier to get, the ARMAX SPM concentration model we present here not only provides a novel method to describe the suspended sediment dynamics but also provides an economical and practical methodology to predict the SPM concentration dynamics based on the current velocity, wave variation and pressure data.

From the simulation results shown in Chapter 6 and Chapter 7, we know that the the multi-layer model presented in the Chapter 7 is better than the one-layer model described in Chapter 6. One of the main reason is that multivariable models take up the more information since the data set contains the ten different layers of information which gives more detail about the sediment profile dynamics. Another reason may be the time scale of the problem since the time scale in Rufiji data is 3.75 minutes and the one in Holderness Coast is one hour. The shorter the time scale, the more efficient is the observed information about the real current velocity and suspended sediment concentration dynamics. More factors such as biological, geophysical information should be considered in the Holderness Coast model if one seeks to improve it.

In this thesis, we consider the current velocity and suspended sediment concentration as stochastic processes which need to be identified. System identification theory is applied in the model which has some unknown parameter matrices to be identified based upon real data collected from the field. It is the principal aim of this thesis to apply the system identification technique to flow and suspended sediment concentration in estuarial and coastal system.

These parameters and parameter matrices will of course be capable of physical, chemical or biological interpretation, but this is not done here since these parameters and parameter matrices are used here only as tools to give a good description of the dynamic system's overall behaviour. We are content to show that this type of model is a good

one in as far as matching known data sets, and we have shown this by showing that all parameters and the elements of each matrix remain virtually constant when subject to future data. This way, the model is shown to describe the data with accuracy.
As we all know, a considerable portion of the sediment transport in the estuary and coast is due to sediment which moves in suspension. It is therefore necessary to develop models for current velocity and suspended sediment concentration dynamics which may in turn be combined with the sediment dynamics to give the transport rate.

Now we turn to discuss and consider further research and future work. It should be pointed out that identification is not a foolproof methodology that can be used without interaction from the user. The next steps we suggest are :

(i) Construct a more appropriate model structure. This can be a difficult problem, in particular if the dynamics of the system are strongly nonlinear.

(ii) There are certainly no 'perfect' data in real life. The fact that the recorded data are disturbed by noise must be taken into consideration.

(iii) The process may vary with time, for example possess natural periodicity or decay, this can obviously cause problems if an attempt is made to describe it with a time-invariant model.

(iv) It may be difficult or impossible to measure some variables/signals that are of central importance for the model since the real cohesive sediment transport dynamics in estuary are still not fully understood.

(v) Consider more variables in the time series model including salinity, temperature, chemical and biological activity.

(vi) Design a three dimensional time series model which includes different stations with their vertical profile data set information and set up dynamical quantitative relationships between the important variables in the sediment transport process.

Chapter 9

Bibliography

Bibliography

[1] Ahmed, M.S. 1986. Parameter Estimation in Bilinear Systems, *Int. J. Control*, **44**, 4, 1177–1183.

[2] Akaike, H., 1970., Statistical predictor identification. *Ann. Inst. Statist. Math.*, **22**, 203-207.

[3] Akaike, H., 1972., Use of an information theoretical quantity for statistical model identification. *Proceeding of the fifth Hawaii International Conference on Systems Science*. Western Periodicals. North Hollywood, California, pp 249–250.

[4] Akaike, H., 1974., A new look at the statistical model identification. *IEEE Trans. Aut. Control*, **19**, 716-722

[5] Akaike, H. 1977., On entropy maximization principle. P.R.Krishniaah (ed). *Applications of Statistics*. Noth-Holland, Amsterdam.

[6] Alexis, A., Bassoullet,P., Le Hir,P. and Teisson,C., 1992. Consolidation of soft marine soils:unifying theories, numerical modelling and in situ experiments. In:*Proc. 23rd Int. Conf. on Coastal Engineering, ICCE'92, Venice*, pp.2949–2961.

[7] Ariathurai, C., 1974. A finite element model for sediment transport in estuaries. *Ph.D thesis*. University of California, Davis.

[8] Åström, K.J., 1980., On the achievable accuracy in identification problems. *1st Prague IFAC Symposium on Identification and Process Parameter Estimation Automatica*, **16**, 281-294

[9] Bagchi, A. and ten Brummelhuis, P., 1996, Parameter Identification in Tidal Models with Uncertain Boundaries. *Automatica*, **Vol. 30, No. 5**, pp.745–759.

[10] Been, K., 1980. Stress strain behaviour of a cohesive soil deposited under water. *Ph.D. Thesis*, Oxford University, 255 pp.

[11] Berlamon, J. Ockenden,M. Toorman, E. and Winterwerp, J.C. 1993. The characterization of cohesive sediment properties. In: H.J. de Vriend (Editor), *Coastal Morphodynamics: Processes and Modelling. Coastal Eng.*, **21**: 105–128.

[12] Billings, S.A. and Voon, W.S.F. 1984. Least Squares Parameter Estimation Algorithms for Nonlinear Systems, *Int. J. Systems Sci.*, **15**, 6, 601–615.

[13] Box, G.P. and Jenkins, G.M. 1976. Time Series Analysis: forecasting and control. *Holden-Day, San Franciso*

[14] Chatfield, C 1980. The Analysis of Time Series : An Introduction *Chapman and Hall, London*

[15] Chen, H.F. 1981a. Strong Consistency of Recursive Identification Under Correlated Noise, *J. Sys. Sci. and Math. Sci.*, **1**, 1, 34–52.

[16] Chen, H.F., 1981b, Quasi Least-Squares Identification and Ite Strong Consistency, *Int. J. Control*, **34**, 5, 921–936.

[17] Chen, H.F. and Guo, L. 1985. Strong Consistency of Parameter Estimates for Discrete-Time Stochastic Systems, *J. Sys. Sci and Math. Sci.*, **5**, 2, 81–93.

[18] Chen, H.F. 1982. Strong Consistency and Convergence Rate of Least Squares, *Scientia Sinica (Series A)*, **25**, 7, 771–784.

[19] Chen, H.F., 1985., Recursive estimation and control for stochastic System. John Wiley and Sons, New York, Chichester, Brisbane, Toronto, Singapore.

[20] Chen, H.F. and Guo, L. 1986. Convergence Rate of Least-Squares Identification and Adaptive Control for Stochastic Systems, *Int. J. Control*, **44**, 5, 1459–1476.

[21] Chen, H. and Dyke, P.P.G. 1995.Time Series Models for Concentration of Suspended Particulate Matter and Its Identification. submitted to *Estuar. Coastal. Mar Sci.*

[22] Chen, H. and Dyke, P.P.G. 1996., Multivariable Time Series Sediment Dynamical Model and Its Identification In Rufiji Delta, Tanzania. *Applied Mathematical Modelling* Vol.20, October, pp.756–770.

[23] Chen, H. and Ruan, R. 1987. Strong Consistency and Convergence Rate of Extended Least Squares Method. *Control-Theory and Advanced Technology*, **3**, No 2. 149-162

[24] Chen, H., Zinober, A.S.I. and Ruan, R. 1996. Strong Consistency and Convergence Rate of Parameter Identification for Bilinear Systems. *International Journal of Control* Vol.63, No.5, pp. 907–919.

[25] Chen,H., Blewett, J., Dyke, P.P.G. and Huntley, D. 1997., Time Series Modelling of Suspended Sediment Concentration on the Holderness Coast. submitted to *Journal of Marine Systems*

[26] Chou, P.Y., 1945., On velocity correlations and the solution of the equations of turbulent fluctuation. *Quart.J.Appl.Math.* 3(1), pp. 38–54.

[27] Chou, Y.C., 1965., Local convergence of martingales and the law of large numbers. *The Annals of Mathematicsal Statistics*, Vol. 36, pp. 552–558.

[28] Chow, J.C. 1972. On estimating the order of an autoregressive moving average process with uncertain observations. *IEEE Trans.Aut.Control*, **17**, 707-709

[29] Cole, P. and Miles, G.V., 1983., Two-Dimensional Model of Mud Transport. *Journal of Hydraulic Engineering Division, ASC.*, Vol.109, No1, January, pp.1–12.

[30] Copeland, A.H., Segall, R.S., Ringo, C.D. and Moore, B.III. 1991. Mathematical modelling of inverse problems for oceans. *Applied Mathematical Modelling*, **15** Nov./Dec. 586-595

[31] Costa, R.G. and Mehta, A.J., 1990. Flow-fine sediment hysteresis in sediment-stratified coastal waters. In:*Proc. 22nd Int. Conf. on Coastal Eng., Delft,* pp. 2047–2060.

[32] Delo, A., 1988. Estuarine muds manual. *HR Wallingford Report* No. SR 164.

[33] De Vantier, B.A. and Narayanaswamy, R. 1989. A suspended sediment flow model for high solids concentration using higher order turbulence closure. *Adavances in Water Resources*, **12**, 46-52

[34] Doob, J.L., 1953.,*Stochastic Process*, John Wiley & Sons, New York.

[35] Dyer, K.R. 1988. Fine sediment particles transport in estuaries: In J. Dronkers and W. van Leussen (Editors), *Physical Processes in Esturaries*. Springer Verlag, Berlin, pp.295–310.

[36] Dyer, K.R. 1989. Sediment Process in Estuarine: Future Research requirements. *Journal of Geophysical Research*, **94**, 14327-14339

[37] Eidsvik, K.J. and Utnes, T. 1991. A model for sediment reentrainment and transport in shallow basin flows. *Coastal Engineering*, **15**, 247-256

[38] Eisma, D. and Kalf, J. 1987. Distribution, organic content and particle size of suspended matter in the North Sea. *Netherlands J. Sea Res.*, **21**, 265-285

[39] Enders, W. 1995 Applied Econometric Time Series *John Wiley & Sons,Inc*

[40] Fennessy, M.J., Dyer, K.R. and Huntley, D.A. 1994. Size and setting Velocity Distribution of Flocs in the Tamar Estuary During a Tidal Cycle. *Netherland Journal of Aquatic Ecology*, **28**, (3-4) 275-282

[41] Fisher, P.R. Characteristic circulation and sedimentation in the Rufiji delta, Tanzania, 1994 *MPhil thesis, University of Plymouth*

[42] Fnaiech, F. and Ljung, L. 1987. Recursive Identification of Bilinear Systems, *Int. J. Control*, **45**, 2, 453–470.

[43] Fritsch, D., Teisson, C. and Manoha, B., 1989. Long-term simulation of suspended sediment transport, application to the Loire estuary. In: *Proc. IAHR Conf., Ottawa*, pp.C276-C284.

[44] Gibson, R.E., Englund, G.L. and Hussey, M.J.L., 1967. The theory of one-dimensional consolidation of saturated clays. *J.Geotechnique*,**17**: 261–273.

[45] Granger, C.W.J., 1980. Forecasting in business and economics. New York. Academic Press.

[46] Green, M.O., Rees, J.M. and N.D. Pearson. 1990., Evidence for the Influence of Wave-Current Interaction in a Tidal Boundary Layer. *Journal of Geophysical Research.* **95**, pp. 9629-9644.

[47] Gust, G., 1976. Observations of turbulent drag reduction in a dilute suspension of clay in sea water. *J. Fluid Mech* 75(**1**): 29–47.

[48] Hagatun, K. and Eidsvik, K.J. 1986. Oscillating turbulence boundary layer with suspended sediments. *Journal of Geophysical Research* ,**91**, 247-256

[49] Hall, P. and Heyde, C.C., 1980., *Martingale Limit Theory and Its Applications*, Academic Press, New York.

[50] Hardisty, J. 1986, Engineering Versus Environmental Requirements in Coastal Prediction. *Journal of Shoreline Management.* Vol. **2**, pp. 65-72.

[51] Harvey, A.C. 1981. Time Series Models. *Philip Allan Publishers Limited*

[52] Hayter, E.J. and Mehta, A.J. 1982., Modelling of Estuarial Fine Sediment Transport for Tracking Pollutant Movement, *UFL/COEL-82/009, Coastal and Oceanographic Engineering Department, University of Florida, Gainesville, Florida.*

[53] Hassain, M.S., 1980. Mathematische Modellierung von turbulenten Auftriebsströmungen, *Ph.D. Thesis*, University of Karlsruhe.

[54] Hsia, T.C., 1977. System Identification. Lexington Books, D.C. Heath and Company, Lexington, Maassachusetts, Toronto.

[55] Huynh-Thanh, S., Hamm, L. and Temperville, A., 1991. Analysis of the erosional behaviour of cohesive sediments with a stratified turbulent transport model. In: *Int. Symposium on the Transport of Suspended Sediments and its Mathematical Modelling, Florence*, pp.423–431.

[56] Jago, C.F., Bale, M.O., Green, M.O., Howarth, M.J., Jones, S.E., McCave, I,N., Millwood, G.E., Morris, A.W., Rowden, A.J., and Williams, J.J., 1993. Resuspension processes and seston dynamics in the Southern North Sea. In. *Understanding the North Sea System.* Edited by H.Charnock, K.R.Dyer, J.M.Huthnance, P.S.Liss,J.H.Simpson,P.B.Tett. Chapman and Hall, London.

[57] Jago, C.F. and Jones, S.E., 1993., Dynamics of Suspended Particulate Matter in the Southern North Sea. I. Tidally Mixed Waters. submitted to *Continental Shelf Research.*

[58] Jones,S., Jago,C.F., Prandle,D. and Flatt,D., 1994., Suspended sediment dynamics: measurement and modelling in the Dover Strait. In. *Mixing and Transport in the Environment.* Edited by, K.J.Beven, P.C.Chatwin and J.H. Millbank. John Wiley and Sons Ltd.

[59] Keller, L. and Friedmann, A., 1924., Differentialgleichungen fur die turbukente Bewegung einer kompressiblen Flussigkeit, *Proc. 1st Int. Congress Appl. Mech., Delft*, pp. 395–405.

[60] Kjerfve, B. and Medeiros, C. Current vanes for measuring tidal currents in estuaries. *Estuarine, Coastal and Shelf Science*, 1989, **28**, 87–93

[61] Krener, A.J. 1975. Bilinear and Nonlinear Realizations of Input-Output Maps, *SIAM. J. Control*, **13**, 4, 827–834.

[62] Krishnappan, B.G., 1991. Modelling of cohesive sediment transport. In: *International Symposium on the Transport of Suspended Sediments and its Mathematical Modelling, Florence*, pp. 433–466.

[63] Krone, P.B., 1962. Flume studies of the transport of sediments in estuarial shoaling processes. Final report, Hyde.Eng.Lab and Sanitary Eng.Res. Lab University of California, Berkeley, CA.

[64] Kujiper, C, Cornelisse, J.M. and Winterwerp, J.C., 1989. Research on erosive properties of cohesive sediments. *J. Geophys. Res,* 94(C10): 14341–14350.

[65] Kynch, G.J., 1952. A theory of sedimentation. *Trans. Faraday Soc.*, **48**: 166–176.

[66] Lee, R.C.K. 1964., *Optimal Estimatioin, Identification and Control* MIT Press, Cambridge, Mass.

[67] Liptser, R.S. and Shiryayev, A.N., 1977., *Statistics of Random Processes*, Springer, New York.

[68] Ljung, L. and Söderström, T. 1983. *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, Mass.

[69] Loeve, M., 1960., *Probability Theory*, Van Nostrand Reinhold.

[70] McGuirk, J.J. and Rodi, W., 1978. A depth-averaged mathematical model for the near field of side discharges into open channel flow. *J.Fluid Mech.*, 86, pp. 761–781.

[71] Mehta, A.J. and Partheniades, E., 1975. An investigation of the depositional properties of flocculated fine sediments *J. Hydraul. Res.,* 13(4): 361-381.

[72] Mehta, A.J., 1986, Estuarine cohesive sediment dynamics. *Lecture Notes on Coastal and Estuarine Studies.* Vol. 14, Springer Verlag.

[73] Mehta, A.J. and Lott, J.W. 1987. Sorting of fine sediment during deposition coastal sediments. In: *Proc. of Conf. on Advances in Understanding of Coastal Sediment Processes,* **1**, New Orleans, LA.

[74] Mehta, A.J., 1988, Laboratory studies on cohesive sediment deposition and erosion. Physical processes in estuaries. *Springer Verlag, Berlin*, pp.427–445.

[75] Mehta, A.J., 1989a, On estuarine cohesive sediment suspension behavior. *J.Geophys. Res.*, 94(C10):14303-14314.

[76] Mehta, A.J., 1989b, Fine sediment stratification in coastal waters. *Proc. Third National Conference on Dock and Harbour Engineering*, Surathkal, pp.487–491.

[77] Mehta, A.J., Hayter,E.J.,Parker,W.R.,Krone,R.B. and Teeter,A.M., 1989. Cohesive sediment transport. I: Process description. *J.Hydraul.Eng.*, 115(8): 1076–1093.

[78] Migniot, C. 1968. Etude des proprietes physiques de defferent sediments tres fins et de leur comportment sous les actions hydrodynamiques. Houille Blanche, 7:591-620.

[79] Migniot, C. 1989. Tessement et rhecologie des vases (part 2). Houille Blanche, 2: 95–111.

[80] Mous, S.L.J. 1993. Identification of the movement of water in unsaturated soils: the problem of identifiability of the model. *Journal of Hydrology*, **143**, 153-167

[81] Ockenden, M.C., Jones, R.J. and Hale, I., 1989. Grangemouth mud properties. Report SR 197, Hydraulics research Ltd. Wallingford.

[82] Ockenden, M.C., 1993. A model for the settling of non uniform cohesive sediment in a laboratory flume and an estuarine field settling. *J.Coastal Res,*, in press.

[83] O'Connor, B.A. and Nicholson, J. 1988. A three-dimensional model of suspended particle sediment transport. *Coastal Engineering*, **12**, 157-174

[84] Odd, N.V.M. and Owen, M.W. 1972., A two-Layer Model for Mud Transport in the Thames Estuary, *Proc. of the Institution of Civil Engineering, London, Supplement* (ix). pp. 175–205.

[85] Parchure, T.M. and Mehta, A.J., 1985. Erosion of solft cohesive sediments deposits. *J. Hydraul.Eng.* 111(**10**):1308-1326.

[86] Perigaud, C. 1983. Mecanique de l'erosion des vases. Houille Blanche, **7/8**: 501-512.

[87] Pringle, Ada W. 1985. Holderness Coast Erosion and the Significance of Ords. *Earth Surface Processes and Landforms.* Vol. **10**. pp. 107-124.

[88] Rastogi, A.K. and Rodi, W., 1978. Predictions of heat and mass transfer in open channels. *Journal of the Hydraulics Division, ASCE*, No. HY3, pp. 397–420.

[89] Rodi, W., 1979. Turbulence models and their application in hydraulics. *IAHR State-of-the-Art Paper*, Delft.

[90] Ross, M.A. and Mehta, A.J., 1989. On the mechanics of lutoclines and fluid mud. *J.Coastal Res., Spec. Iss.*,,5:51–61.

[91] Robbins, H. and Siegmund, D.A. 1971. A Convergence Theorem for Nonnegative Almost Supermartingales and Some Applications. in *Optimizing Method in Statistics*, Academic Press, 233–257

[92] Schröder, M. 1988. Untersuchungen zur Chemie von Schwebstoffen in der Nordsee von Januar bis März. *Diploma thesis, Part 1, Geologisch-Paläontologisches Institut der Universität Hamburg, Germany.*

[93] Sheng, Y.P. and Bulter, H.L. 1982. Modelling coastal currents and sediment transport. *Porceedings of the 18th Coastal Engineering Conference, Cape Town, 14-19 November*, 1127-1148

[94] Sheng, Y.P. and Villaret, C., 1989. Modelling the effect of suspended sediment stratification on bottom exchange processes. *J. Geophys. Res.*, **94(C10)**: 14429–14444.

[95] Simpson, J.H. and Sharples, J. 1991. Dynamically-active Models in the Prediction of Estuarine Stratification. In *coastal and Estuarine Studies* P 101-113. Editor: David Prandle. Dynamics and Exchanges in Estuaries and the Coastal Zone. Springer Verlag.

[96] Smith, T.J. and Kirby, R., 1989. Generation, stabilization and dissipation of layered fine sediment suspensions. *J.Coastal Res., Spec. Iss.*, 5:63–73.

[97] Söderstöm, T. and Stoica, P. ,1989., System Identification *Prentice Hall, Cambridge.*

[98] Suda, N., Kodama, S. and Ikeda, M. 1973., Matrix Theory in Automatical Control, *Japanese Automatical Contrl Association*

[99] Sundermann, J. 1994. Suspended particulate matter in the North Sea : field observations and model simulations. *Understanding the North Sea* Ed. H.Charnock *et al.* 45-52

[100] Teisson, C. and Frisch, D. 1988. Numerical modelling of suspended sediment transport in the Loire estuary. *IAHR Symposium on Mathematical Modelling of Sediment Transport in the Coastal Zone, Copenhagen , 30 May-1 June*, 14-22

[101] Teisson, C. 1991. Cohesive suspended sediment transport: feasibility and limitations of numerical modeling. *J.Hydraul. Res.* 29(6):755-769.

[102] Teisson, C., 1992., Modelling of fluid mud flow and consolidation. *Ph.D. Thesis* Katholieke Universiteit Leuven.

[103] Teisson, C. and Latteux, B. 1986, A Depth-integrated bidimensional model of suspended sediment transport, *Proceedings of the 3rd International Symposium on River Sedimentation ,Jackson*, April,pp. 421-429.

[104] Teisson, C., Simonin, O., Galland, J.C. and Laurence, D., 1992. Turbulence and mud sedimentation. *Proc. 23th Int. Conf. on Coastal Enginering , Venice*

[105] Teisson, C., Ockenden, M., Le Hir, P., Kranenburg, C. and Hamm, L. 1993. Cohesive sediment transport processes. *Coastal Engineering*, **21**, 129–162.

[106] Thomas, W.A. and McAnally, W.H., (1985), User's Manual for the Generalized Computer Program System, Open Channel Flow and Sedimatation, TABS-2,IR HL-85-1, U.S.A.E. Waterways Experiment Station, Vicksburg, MS.

[107] Toorman, E. and Berlamont, J., 1993., Mathematical modelling of cohesive sediment settling and cosolidation. In: A.J. Mehta (Editor), *Nearshore and Estuarine Cohesive sediment Transport* . Coastal and Estuarine Studies **42**. American Geoohys. Union, pp. 167–184.

[108] Tse, E. and Weinert, H.L. 1973., Structure determination and parameter identification for multivariable stochastic linear systems, Paper 20-2, Joint Automatic Control Conference

[109] Utnes, T. 1993., Finite element current and sediment transport modelling. *Continental shelf Research*, **13**, No.8/9. 891-902.

[110] Utnes, T. and Ren, G. 1995., Numerical prediction of turbulent flow around a three-dimensional surface-mounted obstacle. *Appl.Math.Modelling.* Vol.19, pp.7–12.

[111] Van Leussen, W. 1988., Aggregation of particles, settling velocity of mud flocs– a review. In J. Dronkers and W. van Leussen (Editors), *Physical Processes in Esturaries.* Springer Verlag, Berlin, pp.347–403.

[112] Van Leussen, W. and Cornelisse, J.M., 1992. The determination of the sizes and settling velocities of esturine flocs by an underwater video system. *Neth. J. Sea Res.*, submitted.

[113] Van Leussen, W. and Winterwerp, J.C., 1990. Laboratory experimants on sedimentation of fine-grained sediments: A state-of-the-art review in the light of experiments with the Delft Tidal Flume. In: R.T.Cheng(editor), *Residual Currents and Long Term Transport.* Coastal and Estuarine Studies, **38**. Springer Verlag, New York, pp 241–259.

[114] Van Rijn, L.C. 1987. Mathematical modelling of morphological processes in the case of suspended sediment transport. *Dr. Thesis, Technical University Delft, Delft*

[115] Van Rijn, L.C. and Meijer, K. 1988. Three dimensional mathematical modelling of suspended sediment transport in current and waves. *IAHR Symposium on Mathematical Modelling of Sediment Transport in the Coastal Zone, Copenhagen , 30 May-1 June*, 89-99

[116] Veeramachaneni, R. and Hayter, E.J., 1988. Mathematical modelling of sediment transport at tidal inlets. *IAHR Syposium on Mathematical Modelling of Sediment Transport in the Coastal Zone, Copenhagen,* 30 May–1 June, pp. 14–22.

[117] Verbeek, H., Kujiper, C., Cornelisse, J.M. and Winterwerp, J.C., 1993. Deposition of graded natural muds in the Netherlands. In: A.J. Mehta (Editor), *Nearshore and Estuarine Cohesive sediment Transport* . Coastal and Estuarine Studies **42**. American Geoohys. Union, pp.185–204.

[118] Wang, X.F. and Lu, G.Z. 1987. Recursive Identification and Realization Algorithm for a Class of Bilinear Multivariable Systems, *Control Theory and Applications*, **4**, 1, 105–114.

[119] Wang, Z.K., 1965., *Theory of Random Processes (in Chinese)*, Science Press, Beijing.

[120] Wellstead, P.E. 1978., An instrumental product moment test for model order estimation. *Automatica* **14**, 89-91

[121] Winterwerp, J.C., 1989. Flow induced erosion of cohesive beds. A literature survey. Cohesive Sediments Internal Report No. 25 Rijkswaterstaat, Delft Hydraulics.

[122] Wolanski, E.,Chappell, J., Ridd, P. and Vertessy, R., 1988., Fluidization of mud in estuaries. *J. Geophys. Res.*,93(C3):2351–2361.

[123] Woodside, C.M., 1971. , Estimation of the order of linear systems, *Automatica*, **7**, 727-733

[124] Young, P.C., Jakeman, A. and McMartrie, R., 1980. An Instrumental Variable Method for Model Order Identification. *Automatica* 16, 281-294

[125] Young, P.C. and Beven, K.J., 1994. Data-Based Mechanistic Modelling and the Rainfall-flow Non-linearity. *Environmetrics.* Vol.5. 335–363.

[126] Zhang, Y.G. 1983., Identification of a Class of Stochastic Bilinear Systems, *Int.J.Control*, **38**, 5, 991–1002.

[127] Zhao, M.W. and Lu, Y.Z. 1991., Parameter Identification and Convergence Analysis Based on the Least-Squares Method for a Class of Nonlinear Systems, *Int. J. Systems Sci.*, **22**, 1, 33–48.