NONLINEAR IDENTIFICATION AND CONTROL OF MUSCLE RELAXANT DYNAMICS

A thesis submitted to the

UNIVERSITY OF SHEFFIELD

for the degree of

DOCTOR OF PHILOSOPHY

by

Mohammed KHELFA, B. Eng, M. Eng

Department of Control Engineering

University of Sheffield

July 1990.

TO MY MOTHER AND TO THE MEMORY OF MY FATHER

WHO DIED SHORTLY BEFORE

THE ACCOMPLISHEMENT OF THIS

THESIS

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Professor D.A.Linkens, for his excellent guidance and helpful suggestions, encouragement, stimulation and patience throughout the course of this study.

I would like to thank the former student Dr. M. Menad and the research student M. Mhafouf for many stimulating discussions.

I am greatly undebted to the following people (Mourad Fodil Cherif, Mokhtar Bouraada, Said Ramdani, Mehdi Menad, Aoun Djamel, Ahmed Hammadia and his wife Lena and all others for their financial help.

Finally, and most importantly, I must thank my girl friend, Rita for her encouragement and patient.

TABLE OF CONTENTS

SUMMARY

The work reported in this thesis comprised two major parts which are:

1) Off-line nonlinear identification of muscle relaxant dynamics ,

2) Simulation-based design of a variety of controllers (ranging from classical PID to nonlinear self-tuners) for the closed-loop control of muscle relaxation. Relaxant drugs namely, Vecuronium and Atracurium are considered throughout.

Off-line identification studies, using two special nonlinear identification packages (Nonlinear Identification package and Nonlinear Orthogonal Identification package), were carried out to determine nonlinear difference equation models (NARMAX) that best fit (in the least squares sense) recorded data from trials on humans and dogs for each drug. After validation, these models were assumed to represent, in a nonlinear polynomial form, the muscle relaxant drugs pharmacology. Two different approaches were explored for determining the physiological structure of both relaxant drugs:

a) The drug model to comprise a pharmacokinetics part to represent the drug distribution, and pharmacodynamics which are often modelled by using the well known Hill equation.

b) A cross-correlation approach based on Volterra series.

With the relaxant dynamics structure thus fixed, the work proceeded to the control phase. Simple three-term PID controllers were first designed with their parameters being optimised, off-line, using the Simplex method. The non-adaptive nature of this class of controllers makes their robustness open to question when the system parameters for which they have been optimised change. Hence adaptive controllers in the form of linear and nonlinear generalised minimum variance, self-tuners, generalised predictive and nonlinear k-step ahead predictive controllers were also considered. All these latter control approaches are shown to be satisfactory, in terms of transient and steady state performance.

CHAPTER 1

Introduction

Muscle relaxants reduce the muscle tone and form part of anaesthesia. During surgical operations, they enable the surgeon to operate through small incisions and obtain access to deep structures, without interference from muscle action. The contraction of voluntary muscle is caused by the release of the neuromuscular transmitter, acetylcholine. The motor nerve fibre terminals, lying in gutters on the surface of the muscle, form a complex structure known as the motor end-plate. A short-lived depolarization of the motor end-plate membrane by acetylcholine provokes a series of events which leads to the contraction of the muscle.

Muscle relaxants fall into two groups: depolarizing and non-depolarizing agents. Non-depolarizing agents act by occupying the cholinergic receptor sites at the motor end-plate thus preventing depolarization by acetylcholine, whereas depolarizing agents imitate the action of acetylcholine, but with a prolonged duration of action. A number of neuromuscular blocking drugs are available and the ones used in this research are Vecuronium (the experiments were performed on healthy mongrel dogs) and Atracurium (the experiments were performed on a female patient aged 70 years).

The degree of neuromuscular blockade can be measured in terms of either an evoked electromyogram (EMG) or an evoked tension response (Epstein and Epstein, 1973, Ali and Savarese, 1976). This work considers the first of these two methods. This preference is because of the fast EMG response to the electrical stimulation of a peripheral motor nerve (Epstein and Epstein, 1975). The measured (EMG) signal is processed by being amplified, rectified and integrated.

Conventionally, muscle relaxants are administered by the anaesthetist using bolus injections to produce short periods of relaxation. With such manual procedures it is difficult to maintain desired levels of relaxation and over-paralysis

can occur. When this happens, antagonist drugs, such as neostigmine, are administered to assist patient recovery. When these drugs are used in large amounts, post-operative complications can arise. It is therefore desirable to have an automatic control method which would largely relieve the anaesthetist from this task. This is specially desirable for short-acting drugs such as Vecuronium and Atracurium.

Various design and simulation studies have been developed and applied to control muscle relaxation automatically using feedback control principles. Closed-loop control has been clinically shown to be a safe and effective means for continuous drug infusion, and is superior to manual methods (Sheppard et al, 1979, Asbury et al, 1980, Brown et al, 1980, Linkens et al, 1981). The benefits of feedback include the control of system responses to reference inputs, the reduction of the sensitivity of the system response to patient parameter variations, and the reduction of the effects of output disturbances.

Feedback control of continuous infusions of muscle relaxants offers the advantages of precise control of the degree of paralysis and decrease in the undesirable side effects by virtue of the reduction of the total relaxant dosage (Cass et al, 1976 , Linkens et al, 1981).

Computer control of four non-depolarizing agents (Gallamine, d-Tubocurarine, Alcuronium and Pancuronium) was described by Cass et al (1976). The EMG was used as the measurement and the relaxant was administered by a motor-driven syringe pump in the on/off mode rather than by varying the rate of administration. Their method however results in a fairly long induction period because of the need to measure the response to an initial test dose in each case.

Rametti (1985) designed a microcomputer-based control system to induce and then maintain a desired level of muscle relaxation in patients who undergo surgical procedures. The drug d-Tubocurarine was chosen as it is one of the longest acting of

all commonly used relaxants.

Satisfactory regulation with a mean of 74% paralysis for an 80% set-point in human trials using a simple proportional gain feedback controller was achieved by Brown et al, (1980). The offset was removed by introducing an integral action into the control structure using a fixed PI controller. These authors reported occasional oscillations in the closed-loop response.

Linkens et al (1982) used the evoked EMG response and administered the relaxant through an infusion pump. They considered the short acting Vecuronium and the longer acting Pancuronium. Trials were undertaken on both dogs and humans using fixed parameter proportional plus integral control.

Sensitivity studies (Menad, 1984) proved that an optimised three-term (PID) controller, was unable to cope with expected variations in relaxant dynamics. A Smith predictor was later introduced into the overall control structure in an attempt to offset the pure time delay in the relaxant kinetics. This method showed a marked improvement over the fixed PID structure with a significant reduction in the sensitivity of the output response to system parameter changes.

A classical approach (Wellstead et al, 1979), that of pole-assignment, became a popular technique for digital controller design. It is a well-suited method to adaptive control systems. This approach has been opted for by Menad, (1984) for the control of muscle relaxation. Also Linkens (1986) describes various identification, design and simulation studies which have led to successful on-line control of muscle relaxation involving the use of PI, Smith predictor and self-tuning algorithms. An explicit pole-assignment self-tuner approach has been chosen rather than an implicit self-tuner. The reason being that the process dynamics are estimated directly, an important aspect in muscle relaxant investigations since not only is good control required but also a knowledge of relaxant dynamics is likely to have clinical importance.

Denai (1988) applied a self-tuning PID regulator, to muscle relaxation control. The control designs have included a simple proportional plus integral algorithm and a complex self-tuning algorithm based on pole-placement principles. A successful clinical implementation of these algorithms was reported.

Another technique which has been investigated is the knowledge-based fuzzy logic algorithm suitable for heuristic descriptors of both measurment and control (Linkens and Mahfouf, 1988). Using this approach for muscle relaxant drug administration (Pancuronium) showed very successful results.

In contrast to the work done by the above researchers, in this thesis nonlinear identification of muscle relaxant dynamics for both Vecuronium and Atracurorium is discussed. Nonlinear difference equation models (NARMAX) obtained for both the drugs cited above are used to simulate the relaxant system under different control strategies. Any controller design study requires the existence of a mathematical model of the process to be controlled. Muscle relaxation requires a model that is a combination of the drug pharmacokinetics and pharmacodynamics (Hull et al, 1978).

Several compartmental models have been proposed for the pharmacokinetics of muscle relaxants ranging from a one compartment (Gibaldi and Perrier, 1975) to a nine compartment model (Fleischliand and Cohen, 1966). Linkens et al (1981) confirmed that a two compartment model is sufficient to describe drug kinetics for Pancuronium.

This thesis comprises two majors parts:

The first part consists of the use of two extensive identification techniques to estimate nonlinear difference equation (NARMAX) models. A nonlinear identification package (NLI) designed by Billings and co-researchers was used to identify the dynamics of the drug Vecuronium, and another, the nonlinear orthogonal identification package also designed by Billings and co-researchers was used to

identify the dynamics of the drug Atracurium.

The goal of these identification studies is to try to decide on the most appropriate NARMAX models and their subsequent validation. The drug response models should comprises two parts:

1) One part representing linear pharmacokinetics which is obtained by linearizing the NARMAX model and

2) Another part which is often modelled by a nonlinear effect pharmacodynamics in the form of a Hill equation.

An alternative approach is the use of cross-correlation methods based on Volterra series (Billings and Fakhouri, 1982) to explore the underlying structure.

The second part of this thesis consists of the control of muscle relaxation using different strategies. The postulates of control of muscle relaxation are that it be capable of bringing an initially unrelaxed patient to a selectable set-point in as short a time as possible with minimum overshoot, and that it be able to maintain that level within an acceptable band, with possible set-point changes.

An optimized three-term (PID) controller whose parameters are determined using a hill climbing (or simplex) optimization routine (Nelder and Mead, 1965) was used. The application of the general minimum variance self-tuning controller developed by Clarke and Gawthrop (1975,1979) and Clarke et al, (1975) was considered. Another technique falling into the category of self-tuners is the generalized predictive controller (Clarke et al, 1987) based on an explicit algorithm which has showed its robustness when dealing with NARMAX models of both muscle relaxant. Finally two nonlinear adaptive control approaches were considered. The first one is a direct digital control based on a k-step ahead predictor and the second is an extention of the general minimum variance self-tuner.

The work reported in this thesis is concerned with the identification and control of muscle relaxant drugs (Vecuronium and Atracurium) and it is organised

into seven additional chapters which are briefly summarised as follows:

Chapter 2: This chapter opens with an introduction to the background of the mechanism of neuromuscular transmission.

Chapter 3: This gives an overview of system identification of NARMAX models, and the applicability of the various forms of recursive estimation schemes (eg least squares, extended least squares,etc). Also a brief summary of structure detection and model validity tests is given.

Chapter 4: Nonlinear identification of both drugs (Vecuronium and Atracurium) dynamics is considered.

Chapter 5: Pharmacokinetics and Pharmacodynamics of both drugs (Vecuronium and Atracurium) are presented and two methods are introduced:

a) Fitting a physiological structure to the models (where the drug response comprises two parts: a linear pharmacokinetics which transports the drug into the blood and the nonlinear pharmacodynamics represented by a Hill equation).

b) The use of the cross-correlation methods based on Volterra series (Billings and Fakhouri, 1982) for both muscle relaxants.

Chapter 6: This presents the simulation-based control of both muscle relaxants (Vecuronium and Atracurium). Three different strategies are considered to control the NARMAX models:

a) The widely used controller in control system design (three-term or PID controller)

b) Application of the general minimum variance self-tuning controller.

c) A generalized predictive controller

Chapter 7: Nonlinear GMV control of muscle relaxation. Two adaptive control approaches are described

a) A k-step ahead predictor and,

b) The general minimum variance self-tuner to nonlinear systems.

Chapter 8: The major conclusions of this study and recommendations for further research are given.

CHAPTER 2

Neuromuscular Transmission

2.1) Contraction

The function of the muscle is to contract. Muscles are attached to the bones by tendons and act to move these bones with respect to each other. These movements involve three types of activity.

a) Central nervous system, reflex and voluntary activities.

b) Events intervening between the impulse in the motor nerve and the beginning of the contraction.

c) The contractile process itself.

In the present chapter we shall be concerned with the transmission of the impulse from the nerve to the muscle leading to contraction.

The nerve consists of bundles of hundreds of fibres. Each one of which is an extension of a single nerve cell or neurone.

A neurone is a cell with a series of processes. Those processses which bring impulses to the nerve cells are called dendrites.

The cell body gives rise to only one axon which is much longer than the dendrites. The axon which initially can be distinguished from the cell body by the absence of pigment granules and is much less densely covered by synaptic knobs (axon hillock), leaves the cell body and makes contact with an effector organ, in this case a number of muscle fibres.

The axon of one neurone either forms a synapse with dendrites or arborises directly around the cell body. Each axon lies in a hollow tube (neurilemma) beneath a sheath called endoneurium which is constricted at intervals to form the node of Ranvier. Each segment of peripheral nerve contains a nucleated cell (shwann cell).

Medullated nerves have sheaths of myelin between axon and neurilemma. A schematic representation of a neuron is shown in Figure 2.1(a).

The structure of the synapse is diagramatically represented in Figure 2.1(b), which shows the relationship of the axon terminals to the dendrites of the cell body.

The details of the synaptic function are shown in Figure 2.2 and discussed in section 2.2.

The contraction is led by a number of events which occur in the synapse between the nerve and the muscle, the so-called neuromuscular junction. These events are shown in Figure 2.3.

2.2) Anatomy of the neuromuscular junction.

The neuromuscular junction is a long cell containing many nuclei with fibrils embeded in a matrix of sarcoplasm, the whole being surrounded by a membrane called sarcolemna. The junction between the fibre and its motor nerve is known as the end-plate (Gutman and Young, 1944). A diagram of the structure is illustrated in Figure 2.2. According to other studies (Couteaux, 1944) the end-plate has a mixed embryonic origin, the tissue adjacent to the nerve fibrils being derived from shwann sheaths. The use of modern techniques (Birks et al, 1960) shows more details of the structure. The motor nerve is separated from the muscle by the synaptic cleft. The structure of this cleft is as follows. The motor nerve ends at the part of the muscle membrane known as the motor end-plate. In the area of the end-plate, the membrane lies in gutters on the surface of the muscle to constitute the synaptic cleft. At this level under the ridge of each gutter orifices to secondary cleft or cholinesterase can be found.

A brief schematic of the structure of the neuromuscular junction is shown in Figure 2.1(b).

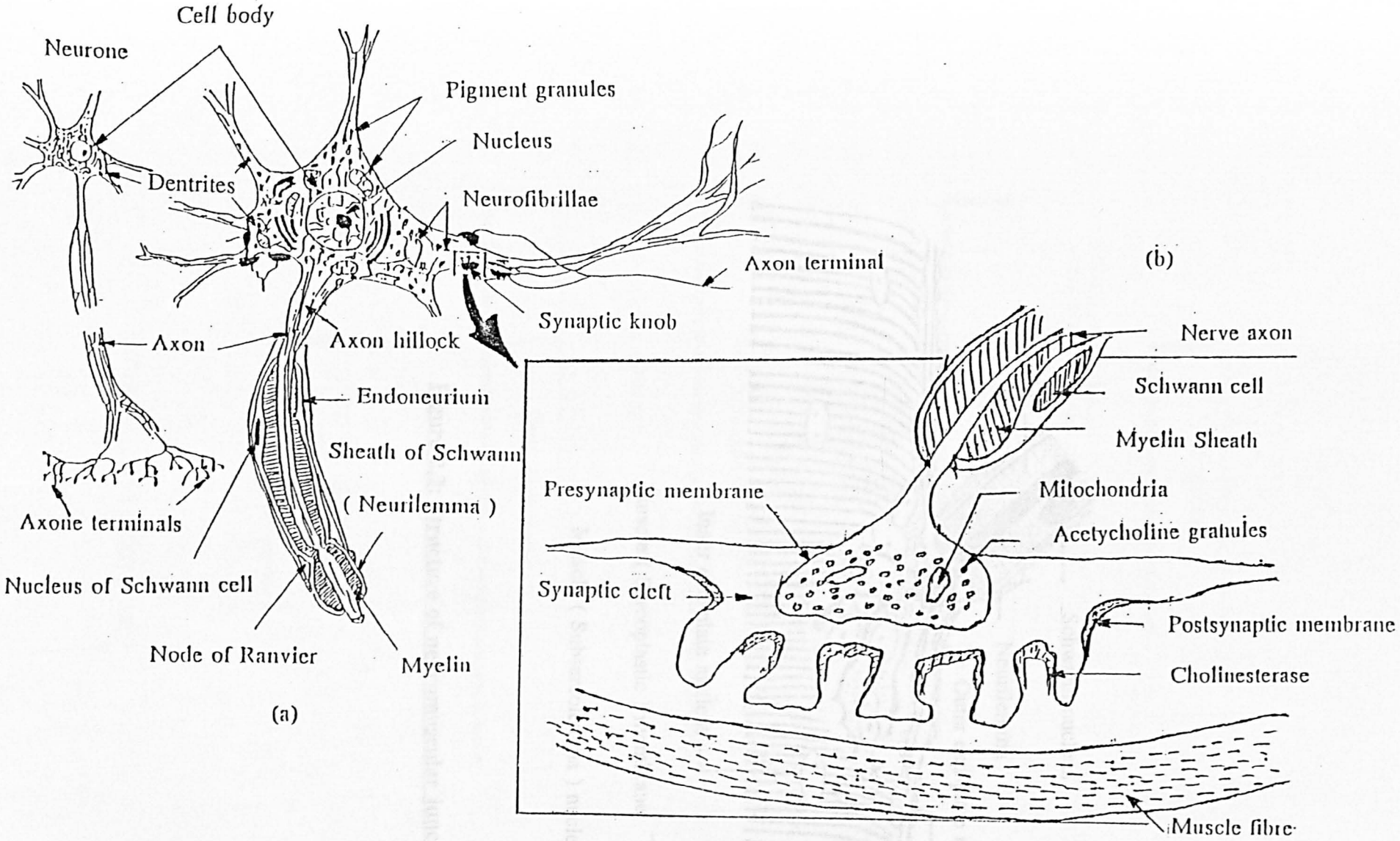


Figure 2.1: A schematic representation of a neurone and the structure of the synapse.

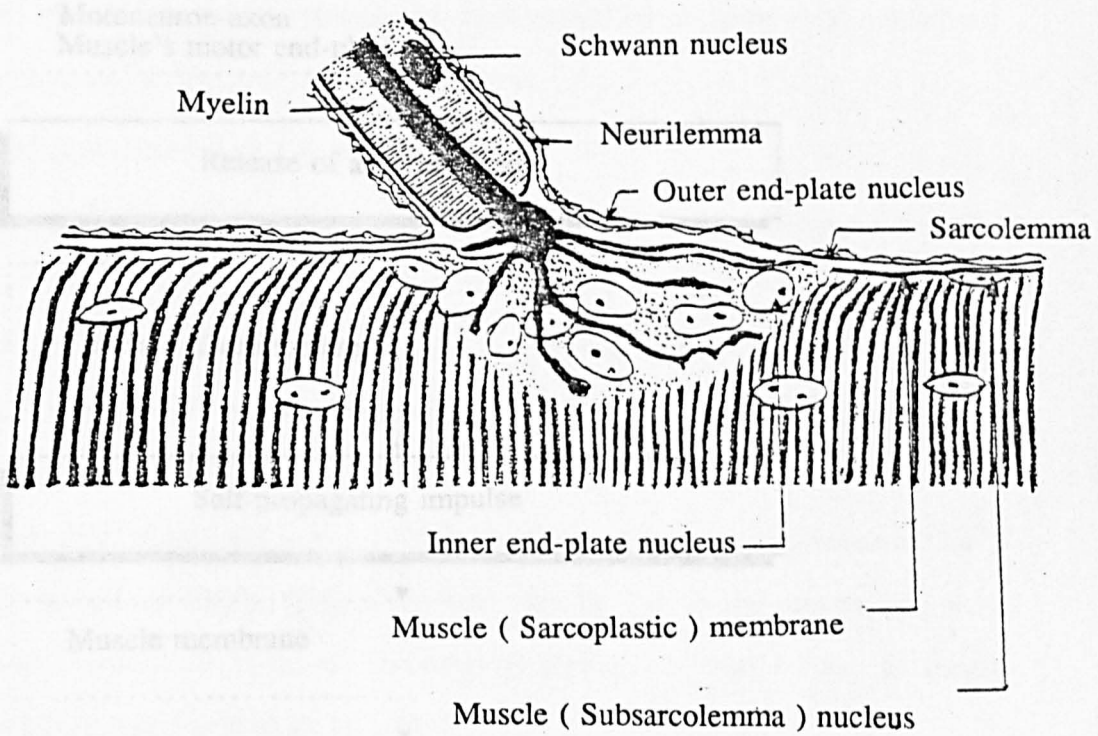


Figure 2.2: Structure of neuromuscular junction

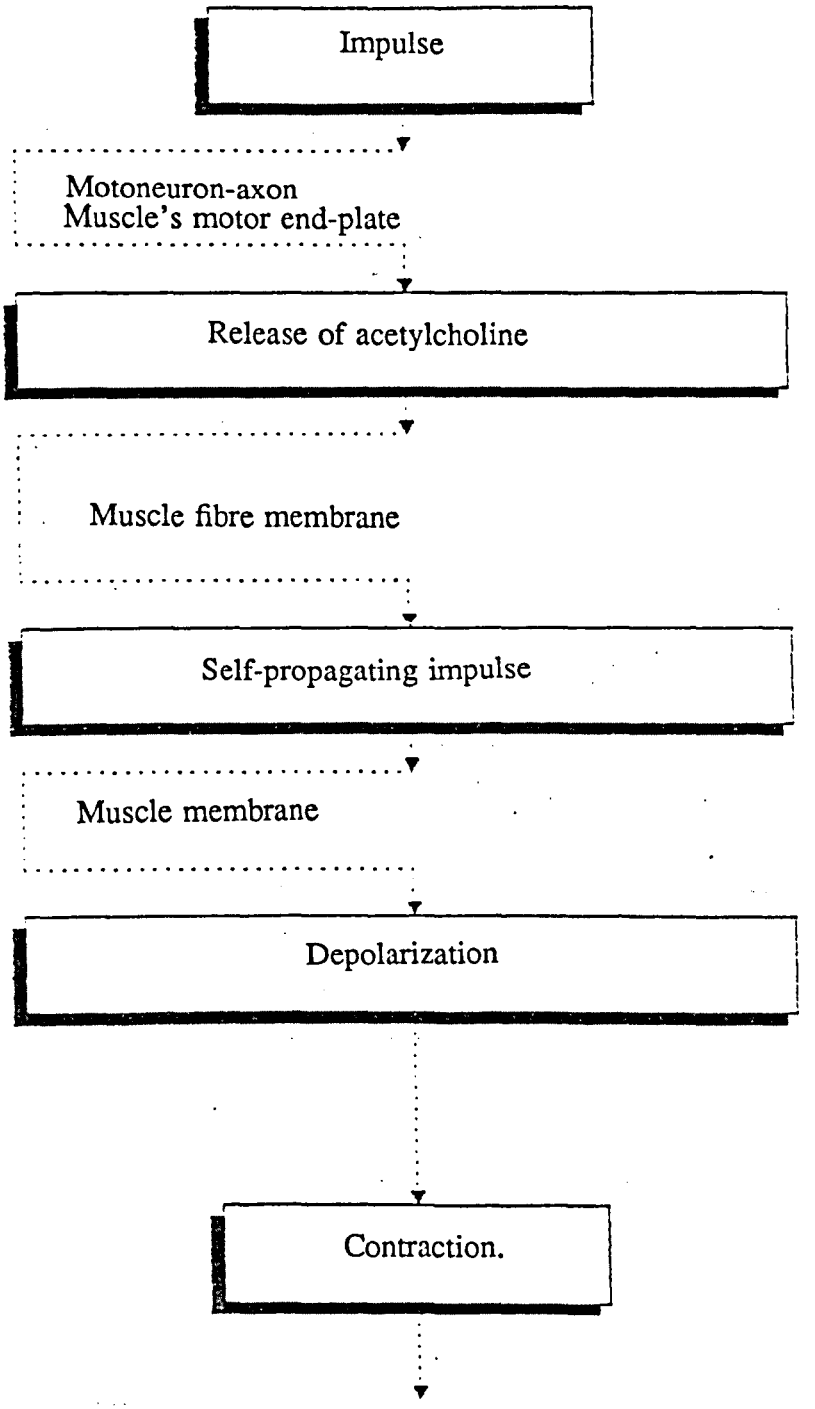


Figure 2.3: Contraction events

2.3) Neuromuscular transmission.

Woodbury et al (1965) and Birks et al (1960) showed, with physiological evidence and consideration of the structure of the nerve terminal and end-plate membrane, that the neuromuscular transmission is accomplished by other means than local circuit flow. This transmission is mediated by a chemical phenomenon, due to a chemical substance called acetylcholine.

2.3.1) Polarization of the membrane.

At rest, the inside of the motor end-plate membrane has a negative and the outside a positive electrical charge. This means, that the membrane has a resting potential which can be measured (Graham and Gerard, 1946). The membrane has a porous structure. These pores are of such a size that ions can move through them rapidly.

Also, at rest, the membrane is about 100 times more permeable to potassium (Ka^{+}) than it is to sodium (Na^{+}). This difference may be due to the greater radius of (Na^{+}) ions. This leads to an electrochemical resting potential across the membrane which is hence said to be polarized.

2.3.2) Depolarization of the membrane.

The arrival of an impulse at the nerve terminals releases an amount of acetylcholine to cause depolarization of the post-synaptic membrane by increasing its permeability to (Na^{+}) ions. The membrane becomes more permeable to sodium than potassium. The membrane potential is therefore reversed by this chemical phenomenon.

These depolarization chemical sequences lead to the depolarization of the adjacent electrically excitable muscle membrane (Waud, 1968) which gives rise to an action potential followed by muscle contraction. But this depolarization is short lived, due to the presence of the enzyme acetylcholinesterase in high concentrations in the end-plate regions of the membrane.

2.4) Blockage of the neuromuscular transmission.

By their mechanisms of action, muscle relaxant agents fall largely into two categories: The depolarisers and the non-depolarisers.

Thus far (section 2.3), acetylcholine has been described as a reactive substance as far as the junction region is concerned. It has been shown (Rike, 1975), that the neuromuscular blocking agents affect the release and production of acetylcholine. Drugs such as succinylcholine, sccinylcholine and decamethonium, known as depolarisers, have a similar action to that of acetylcholine at the post-synaptic membrane.

Among the depolarisers mentioned above, the most commonly used is the succinylcholine drug. It is the 'best' short-acting neuromuscular blocking agent (Savarese and Kitz, 1975) and an excellent clinical drug, safe in most situations.

2.4.1) Depolarisers and non-depolarisers.

The depolarisers facilitate an increase in the production of acetylcholine and activate the cholinergic receptors, which lead to depolarisation. The depolarization persists inspite of further release of acetylcholine to excite the muscle and paralysis will follow.

The non-depolarisers which include Dimethyltubocurarine, Gallamine, Pancuronium, Vecuronium, Atracurium, Alcuronium, D-tubocurarine prevent depolarization of the end-plate hence blocking the neuromuscular transmission. These agents also enter in competition for the cholinergic receptors with acetylcholine hence the name "competitive block".

The reaction of acetylcholine with cholino-receptors leads to the failure of sufficient sodium pores to open, to permit a threshold depolarisation to take place. This activation causes a disturbance at the surface of the membrane and hastens the displacement of non-depolarising drugs. If after establishing a neuromuscular block, the environmental concentration is reduced rapidly, the non-depolarisers can

be shown to exhibit different dissociation rates from the receptor. This is known as their recovery index.

2.4.2) Desensitization.

The action of the drug succinylcholine, mentioned in section 2.4, is short lived. The use of succinylcholine, with time and high total dosage, will produce two types of block. The initial block is depolarising (phase I type) and the late form is non-depolarising and is referred to as "desensitising block" (phase II, dual or mixed block).

Katz (1974) believes that the assumption that the prolonged response to succinylcholine is due to changes in the nature of the block from depolarising to desensitizing is incorrect. The much more likely explanation given by Katz is overdose.

2.4.3) Margin of safety.

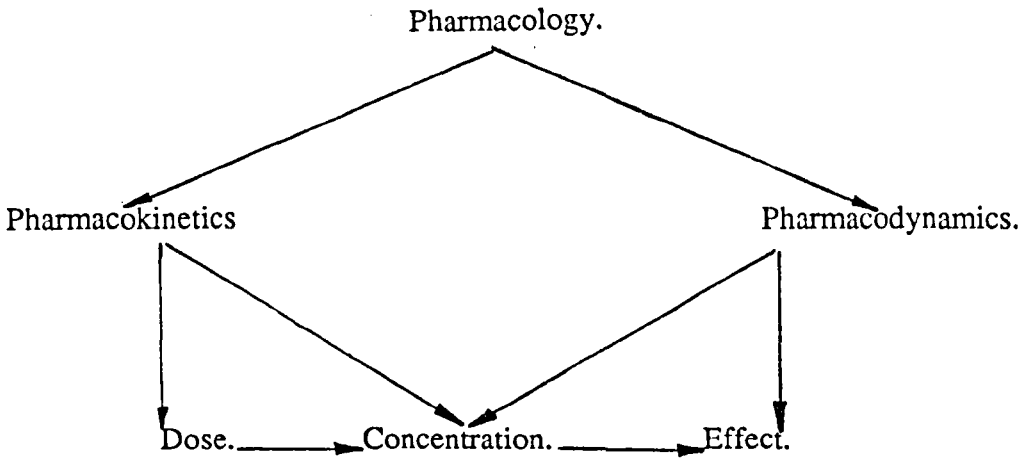
Paton and Waud (1967) introduced the concept of the margin of safety of neuromuscular transmission, which was later developed (Waud and Waud, 1971). The excess of acetylcholine released during a motor nerve activity will produce critical depolarisation of the post-synaptic membrane. However, in the presence of non-depolarising relaxants, the change in membrane potential is not sufficient to trigger off a propagated action potential. Hence the presence of "margin of safety" of neuromuscular transmission.

2.5) The modelling of pharmocological effects of relaxant drugs.

The concept of pharmocology is the relation between the administration of a drug and the pharmacological effect it produces. This concept consists of two parts: pharmacokinetics and pharmacodynamics.

The pharmacokinetics refers to the rate of change in drug concentration within the body and its component organs, tissues and fluids. The practice of pharmacokinetics describes the relationships between drug dose and drug concentration in the blood, plasma or other body components. The kinetics of drug disposition cannot

be isolated from the pharmacodynamics. The pharmacodynamics describe the relationship between the drug concentration and response. All these terms can be related as follows (Hug and Roberts, 1984).



2.5.1) Pharmacokinetics.

Pharmacokinetics modelling has been considered by many authors (Ham et al, 1979), (Stanski et al, 1979) and (Hull et al, 1980). In biological systems, models are often formulated in terms of compartments. Compartmental models have been widely adopted in biomedical research. Organs or other entities are connected by various pathways. Movements of a drug from compartment to compartment within the model are represented as first order processes.

a) One compartment model.

Relaxant drugs are rapidly distributed into the plasma and other body tissues and fluids upon entry into the system circulation following an intravenous (iv) injection as shown in Figure 2.4.

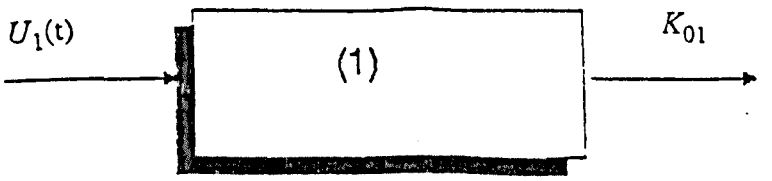


Figure 2.4: A one compartment model

This is modelled as a single exponential decay of concentration in the blood (Gibaldi and Perrier, 1975) and (Jacquez, 1972).

The differential equation describing the model is:

$$\frac{dx_1}{dt} = -k_{01}x_1(t) + u_1(t) \tag{2.1}$$

where

$u_1(t)$ is the drug input

k_{01} is the transfer rate from the compartment to the environment

$x_1(t)$ is the total amount of drug in the body at time t

For a rapid intravenous injection and if perturbation about a steady state is being considered we have:

$$x_1(t) = u_1 e^{-k_{01}t} \tag{2.2}$$

b) Two-compartment models.

A two compartment model for drug distribution is shown in Figure 2.5

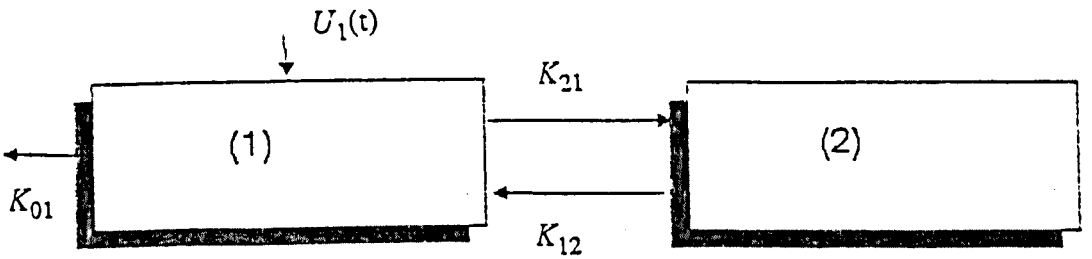


Figure 2.5: A two-compartment model

Shanks et al (1980) suggested two compartment models for both d-Tubocurarine and Pancuronium. The compartment (1) is the central compartment representing the highly perfused tissues (heart, kidneys, liver and lungs) while compartment (2)

is the peripheral compartment representing the poorly perfused tissues (muscle, skin and fat).

k_{12} denotes the transfer rate from compartment (1) to compartment (2), the forward transfer.

k_{21} denotes the transfer rate from compartment (2) to compartment (1), the backward transfer.

The plasma concentration (C) takes the following bi-exponential form following an intravenous injection of a relaxant agent.

$$C(t) = ae^{-ct} + be^{-dt} \tag{2.3}$$

where

a, b, c and d are constants.

c) Three compartment models.

The interpretation of the three compartment models was studied by several authors : (Brown and Godfrey, 1978) for the metabolism of bilirubin, (Sumner et al, 1976) for dioxin and (Gibaldi et al, 1972) for d-Tubocurarine in man.

Gibaldi et al (1972) proposed an appropriate pharmacokinetics model for d-Tubocurarine in man. It comprises three compartments Figure 2.6

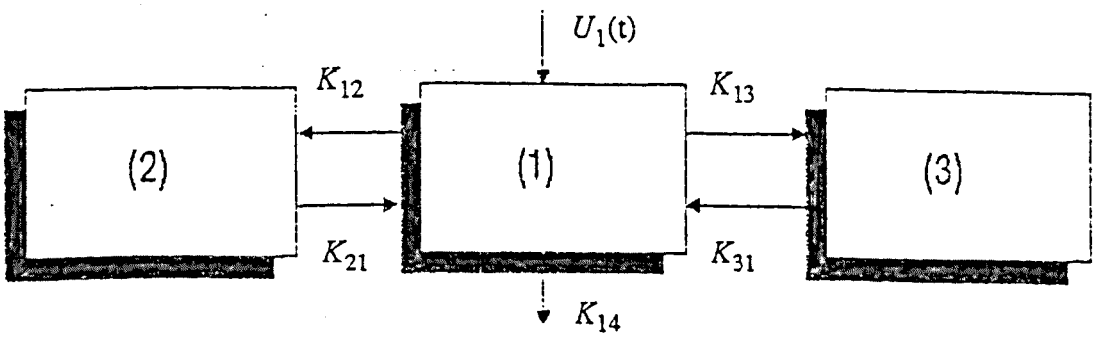


Figure 2.6: A three compartment model

The central compartment (1) corresponds to the plasma, where the relaxant drug is

injected and excreted into the urine. The peripheral compartments (2) and (3) represent other sites unaccounted for. Frequently compartments (2) and (3) are interpreted as extracellular and intracellular compartment of poorly perfused tissues. The impulse response of the plasma concentration in the compartment of interest is a sum of decaying exponentials at time t following an intravenous injection of d-Tubocurarine

$$x(t) = ae^{-bt} + ce^{-dt} + fe^{-gt} \tag{2.4}$$

where

a, b, c, d, f and g are constants.

2.5.2) Pharmacodynamics

As mentioned previously, the work of some authors, in biology and medicine, in modelling the pharmacokinetics of drugs was based on compartmental analysis. Regarding pharmacodynamics, Ham et al (1979), Sheiner et al (1979), Wagner (1968) and Whiting and Kelman (1980) extented the classical pharmacokinetics model by modelling a defined effect compartment linked to the plasma compartment by a first order process in such a way that this modification does not affect the parameters of the original model (Figure 2.7).

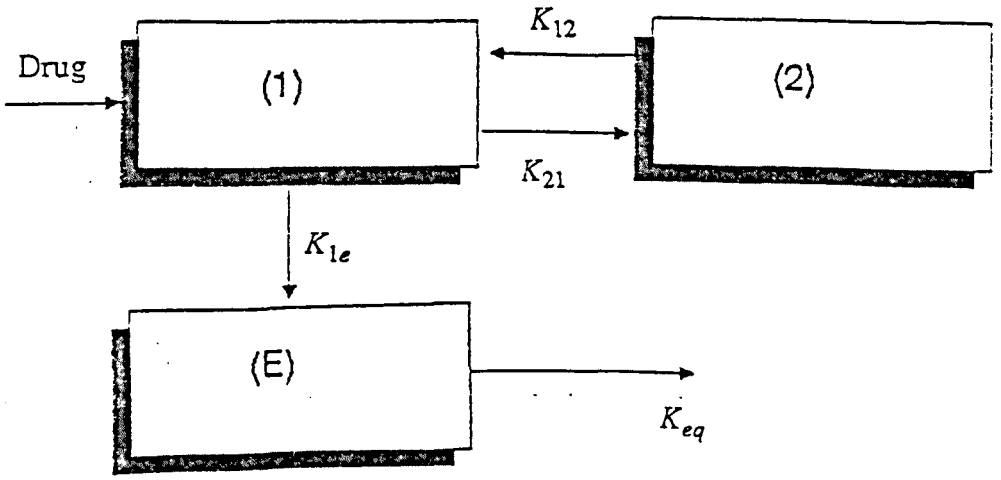


Figure 2.7: A two compartment model with effect compartment

Where

E is the effect compartment

k_{le} a rate constant of very small numerical value so that the original pharmacokinetics model is not affected.

k_{eq} is the plasma concentration effect .

The response is described as a function of the drug concentration C_e:

$$E = f(C_e) \tag{2.5}$$

The function f can be a sigmoid as defined by a Hill equation.

E is thus expressed as

$$E = \frac{E_m C^{\alpha}}{C^{\alpha} + (C_{(50)})^{\alpha}} \tag{2.6}$$

where

E is the drug effect

$C_{(50)}$ the drug concentration in the blood plasma required to produce 50% of Em.

the exponent, α, governs how rapidly the response will approach its maximum as C increases.

It is also worth mentioning that equation 2.6 is transformable into a logistic function.

Consider the general mathematical form of equation 2.6

$$Y_h = \frac{X^{\alpha}}{X^{\alpha} + D^{\alpha}} \tag{2.7}$$

where

Y_h represents the pharmacodynamic effect. Equation 2.7 can be written as:

$$Y_h = \frac{1}{1 + \left[\frac{D}{X}\right]^{\alpha}} \tag{2.8}$$

setting

$$\left[\frac{D}{X}\right]^{\alpha} = e^{-X'} \tag{2.9}$$

therefore

$$X' = \alpha LogX - \alpha LogD \tag{2.10}$$

$$e^{X'} = X^{\alpha}D^{-\alpha} \tag{2.11}$$

hence

$$e^{-X'} = \left[\frac{D}{X}\right]^{\alpha} \tag{2.12}$$

Substituting equation 2.12 into 2.8

$$Y_L = \frac{1}{1 + e^{-X'}}; \qquad (X \geq D) \tag{2.13}$$

This is called the logistic equation. Ham et al, (1979) obtained the complete pharmacological response using three compartment models and the logistic equation. Both expressions of equations 2.7 and 2.13 are non-linear sigmoidal forms.

2.6) Monitoring, types of stimulants and reversal of neuromuscular blockade.

Whenever muscle relaxants are administered it is necessary to monitor the neuromuscular transmission to provide a guide to dosage. The most important aspect of monitoring is the determination of reversal of non-depolarizing blocks. Many authors developed different types of stimulation (twitch, tetanus, train-of-four and electromyograph).

(Feldman, 1987) found that the electromyographic methods are suitable for use in small animals as they require less apparatus. But the disadvantage of electromyography is that it is difficult to obtain good records during clinical conditions in the operating theatre. Other researchers assessed the response using the twitch

method (Epstein and epstein, 1973). Others favour the train-of-four methods (Ali et al, 1970, 1975).

In spite of the requirement for more sophisticated techniques the electromyograph is likely to remain the most widely used in the foreseeable future.

An electromyographic signal can be rectified, integrated, and amplified to produce a voltage proportional to the area under the response curve.

2.7) Administration of muscle relaxant drugs.

Administration of muscle relaxant drugs is usually carried out manually by anaesthetists. If prolonged paralysis is needed for adequate operating conditions, supplements of doses are administered. This might result in over-paralysis at the end of an operation. When over-paralysis occurs, antagonist drugs such as atropine and neostigmine are administered to reverse the neuromuscular block. But this anticholinesterase can cause severe cardiovascular problems.

2.8) Interaction of relaxant drugs with disease and other drugs.

In this section a brief summary is given of the interaction of relaxant drugs with some diseases and other drugs (Aarow et al, 1973).

Diseases such as myasthenia gravis and carcinomatous neuromiopathy cause neuromuscular disfunction. In the case of the myasthenic patients, the basic pathophysiology is not known but reports suggest that disease is due to the reduction in the quantal amount of acetylcholine from the motor nerve terminals. To achieve muscle relaxation, these patients need large doses, which rapidly develop to prolonged desensitizing-type blocks. To overcome this problem, anaesthesia is best done using inhalant anaesthetics.

Early in anaesthesia, if administration of d-tubocurarine follows recovery from the use of succimylcholine, it interacts with the desensitising phase. If the administration of drugs is reversed (for rapid relaxation), this leads to the depolarizing action of succinylcholine which antagonizes the effect of d-tubocurarine and tends to

reduce neuromuscular relaxation.

Antibiotics influence neuromuscular transmission (calcium, will usually antagonize the neuromuscular blocking effect). Also simultaneous use of relaxants and antibiotics will possibly disturb the respiratory function.

Many drugs and chemicals such as insecticides and some of the antimetabolites reduce the plasma cholinesterases.

Simultaneous use of relaxants and cardiac drugs which include procaine, procainamide, quinine, quinidine, dilantin and lidocaine may severely depress neuromuscular transmission.

CHAPTER 3

Nonlinear System Identification

3.1) Introduction

Over the years, the mathematical modelling of processes required knowledge of physics, chemistry, biology etc to develop the equations governing their dynamics. Because of the complexity of physical phenomena, in some cases the laws of science are inadequate to give a satisfactory discription of the plant dynamics. However, researchers used other means of information about the dynamics, which are the data obtained from experiments. The process of finding a model and estimating the unknown parameters from this data is called system identification (Goodwin et al, 1977).

The aim in many identification problems, in both industry and the biosciences is to combine a priori knowledge with experimental data. System identification is a well established subject in control theory. Although this discipline is well developed for linear systems, the identification of nonlinear systems is not well exposed.

Every physical system is nonlinear in some way. It is essential that more effort is directed towards nonlinear system identification.

This chapter presents a brief summary of three distinct types of models for the presentation of nonlinear systems, an introduction to the nonlinear identification package (NLI), a summary of the algorithms it employs and an illustrative example of its use, an introduction to the nonlinear orthogonal identification (NOI), and the last section deals with a brief summary of other packages used in this research.

3.2) Nonlinear models

There are basically three distinct approaches to nonlinear system representation (Billings, 1980 and 1985).

3.2.1) Functional series models

Examples are Volterra series and Wiener series. The Volterra method is used to represent a very broad class of nonlinear systems (e.g communication theory). The concept of this method can be illustrated as follows.

Consider the nonlinear system shown in Figure 3.1.

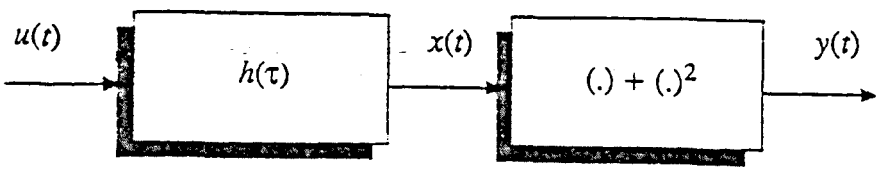


Figure 3.1

Where

$h(\tau)$ is the impulse response of the linear block.

x(t) is the output of the linear block given by the convolution integral.

$$x(t) = \int_0^\infty h(\tau)u(t-\tau) \tag{3.1}$$

the output y(t) of the nonlinear block is given by

$$y(t) = [x(t)] + [x(t)]^2$$

$$= \int_0^\infty h(\tau_1)u(t-\tau_1)d\tau_1 + \int_0^\infty\int_0^\infty h(\tau_2)h(\tau_3)u(t-\tau_2)u(t-\tau_3)d\tau_2 d\tau_3 \tag{3.2}$$

and therefore the general Volterra representation is of the form (Volterra,1930)

$$y(t) = \sum_{n=1}^{\infty}\int_0^\infty \cdots \int_0^\infty h_n(\tau_1,\tau_2,\ldots,\tau_n)\prod_{i=1}^{n}u(t-\tau_i)d\tau_i \tag{3.3}$$

where u(t), y(t) are the input and output signal respectively and the function $h_n(\tau_1,\tau_2,.....,\tau_n)$ is the n' th order Volterra kernel.

The identification of nonlinear systems based on the Volterra representation requires estimation of the Volterra Kernals $h_n(\tau_1,\tau_2,.....,\tau_n)$.

Several approaches are used to solve this problem, the most common being the correlation methods used for linear systems.

An example described by Volterra series with just two terms takes the form.

$$y(t) = \int_0^\infty h_1(\tau_1)u(t-\tau_1)d\tau_1 + \int_0^\infty\int_0^\infty h_2(\tau_1,\tau_2)u(t-\tau_1)u(t-\tau_2)d\tau_1\tau_2 \qquad (3.4)$$

and was considered by Billings (1985).

Using cross-correlation methods the solution was extremely difficult for a general stochastic input. If however the input is selected to be a white gaussian noise, the estimation of the Kernal coefficients, becomes straight forward. Unfortunately, in practice, the number of terms in the Volterra series will be unknown and the above method would break down whenever higher terms should be present. The functional series representation has been studied by several authors (Fakhouri, 1980), (Korenberg, 1982), (Schetzen, 1980) and (Marmarelis et al, 1978). However little progress was achieved in terms of practical applications.

3.2.2) Block-orientated models

A general representation of the structure of block orientated models, analysed by many authors (Billings and Fakhouri, 1978 and Korenberg, 1982) is depicted in the cascade system of Figure 3.2.

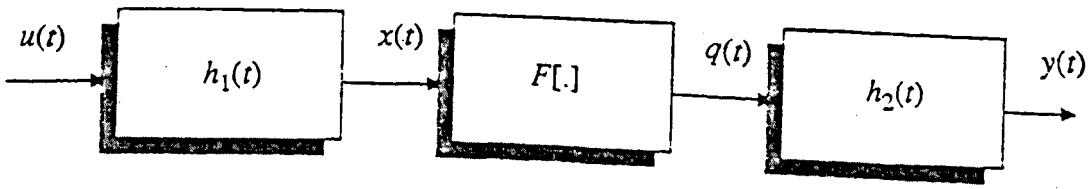


Figure 3.2

where

x(t) and q(t) are internal signals, which, realistically, are unavailable for measurement.

(Billings et al, 1982) utilised the method of separability. Other authors (Haber et al, 1976) have considered sub-classes of the general model referred to as the Wiener and Hammerstein models represented in Figure 3.3.

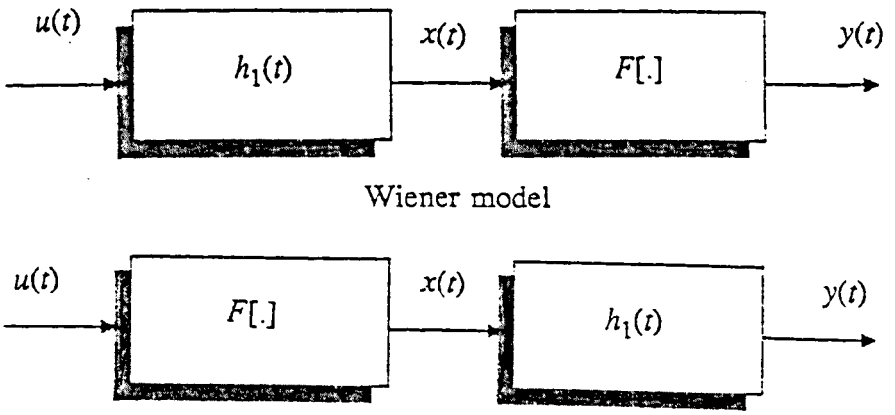


Figure 3.3

It is clear from Figures 3.2 and 3.3 that the solution of the identification problem is to estimate $h_1(\tau)$, $h_2(\tau)$ and F[.] (Billings et al, 1982). The main advantage of the block orientated approach is that it is easy to identify and to relate the results to the physical system. Even systems with severe nonlinearities can be identified using this approach.

3.2.3) Difference equation models

The most important class of models in identification and digital control are based on linear differential equations (Goodwin et al, 1977), (Goodwin et al, 1984) and (Iserman, 1972).

The representation of nonlinear models by nonlinear differential equations has a

prominent role in identification because the development of the model from physical laws arises naturally, directly or after some approximation. Usually it contains few parameters and is therefore easy to handle numerically. The coefficients of these models can be estimated using parameter estimation algorithms.

If an approximate linear difference equation model is fitted to nonlinear data, this will provide a poor fit (Billings et al, 1985). It is preferable to use a nonlinear model than an approximate linear model. If the system is linear, then it can be represented by the following difference equation (Goodwin et al, 1977)

$$y(t) = -\sum_{i=1}^{n_y}[a_i y(t-i)] + \sum_{i=1}^{n_u}[b_i u(t-d-i)] \tag{3.5}$$

where

y(t) is output sequence at time t

u(t) is the input sequence at time t

d is the system time delay

and

a_i , b_i are the unknown model parameters. All the parameter estimation algorithms for linear systems are based on this model which provides a very concise system representation.

(Billings et Leontarities, 1981, 1982) showed that a nonlinear system can be represented by a nonlinear difference equation of the form

$$y(t) = F_*[y(t-1),...,y(t-n_y);u(t-d),...,u(t-d-n_u+1)] \tag{3.6}$$

where

F_* [.] is some nonlinear function of u and y.

The model of equation 3.5 is usually referred to as an Auto-Regressive Moving Average (ARMAX) model. And the model of equation 3.6 will be referred to as the nonlinear Auto-Regressive Moving Average with eXogenous inputs or (NARMAX) model.

A general representation for a wide class of nonlinear systems is of the form

$$z(t) = F_*^l[z(t-1),...,z(t-n_y);u(t-d),...,u(t-d-n_u+1);e(t-1),...,e(t-n_y),e(t)] \quad (3.7)$$

For more details see appendix A.

3.3) Structure detection and model validity tests

In the early stages of an identification study it is important to determine if the process under test exhibits nonlinear characteristics which will warrant a nonlinear model. If the data obtained is from a linear system, a linear model is considered, otherwise a nonlinear model ought to be considered.

3.3.1) Structure detection test (STD)

STD's are used to test whether the system under investigation is linear or nonlinear. Fadzil (1985) offers several methods : step tests, output mean tests, Subba Rao and Gabr tests and correlation tests.

The correlation test is a very useful method. The computation required is not excesssive and interpretation of the results is straight-forward.

The study of the structure detection tests based on correlation analysis has shown that when the system is nonlinear all the third order moments of the input are zero and all even orders moments exist, thus:

$$\phi_{y'y'^2}(\tau) = E[y'(k+\tau)(y'(k))^2] \quad (3.8)$$

where

y' indicates that the mean level has been removed.

If the process is linear then

$$\phi_{y'y'^2}(\tau) = 0, \text{ for } all\ \tau \quad (3.9)$$

If

$$\phi_{y'y'^2}(\tau) \neq 0, \text{ for } all\ \tau \quad (3.10)$$

then the system under test is nonlinear (Billings et al, 1983).

3.3.2) Model validity test (MVT)

Several methods have been developed for MVT and the most useful one is based on correlation analysis techniques. Many studies (Billings and Voon, 1983)

have indicated that fitting linear models to data with significant nonlinearities can provide very misleading results. This occurs, because the linear parameter routines tend to yield estimated combination of process and noise model, which visually provides a good prediction of the system output. The model is however usually significantly biased.

The MVT can fail to indicate the inadequacy of the estimated model when the process is nonlinear. This can mislead the experimenter into believing that the model is adequate when it is not. If the system is linear, the residuals should be unpredictable from all past inputs and outputs.

When the system is nonlinear, the residuals should be unpredictable from all linear and nonlinear combinations of past inputs and outputs and this condition will hold if and only if

$$\left.\begin{aligned} &\phi_{\xi\xi}(\tau) = \delta(\tau) \\ &\phi_{\xi\xi}(\tau) = 0 \text{ for } \textit{all } \tau \\ &\phi_{\xi\xi u}(\tau) = E[\xi(t)\xi(t-1-\tau)u(t-1-\tau) = 0, \text{ for } \textit{all } \tau \end{aligned}\right\} \qquad (3.11)$$

where

$\phi_{\xi\xi}(\tau) = E[\xi(t)\xi(t-\tau)]$ detects deficiencies in the noise model.

$\phi_{\xi u}(\tau) = E[\xi(t)u(t-\tau)]$ detects all odd terms in f(u(t),y(t)).

$\phi_{\xi\xi u}(\tau) = E[\xi(t)\xi(t-1-\tau)u(t-1-\tau)]$ detects terms of the form $u^q(t-m)e(t-m)$, for all m,n odd q.

For nonlinear systems, the tests $\phi_{\xi\xi}(\tau)$ and $\phi_{\xi u}(\tau)$ are not sufficient.

If instrumental variables or suboptimal least-squares are used, the residuals may be coloured.

It can be shown that in this case the process model is unbiased if and only if

$$\left.\begin{aligned} &\phi_{u\xi}(\tau) = 0, \text{ for } \textit{all } \tau \\ &\phi_{u'^2\xi}(\tau) = 0, \text{ for } \textit{all } \tau \\ &\phi_{u'^2\xi^2}(\tau) = 0, \text{ for } \textit{all } \tau \end{aligned}\right\} \qquad (3.12)$$

where

$\phi_{u'^2\xi}(\tau) = E[u'^2(t)\xi(t+\tau)]$ detects all even terms in f(u(t),y(t))

$\phi_{u'^2\xi^2}(\tau) = E[u'^2(t)\xi^2(t+\tau)]$ detects all missing terms e(t) and power of e(t).

Equations 3.11 and 3.12 give the experimenter a great deal of information regarding the deficiencies in the fitted model and can indicate which terms should be included in the model to improve the fit (Billings and Voon, 1983).

3.4) Nonlinear identification package (NLI) for nonlinear systems

The NLI package was originally written by Voon and Leontaritis. It has been modified by Fadzil to constitute a powerful interactive package consisting of several suites of programs for data transformation, structure detection, parameter estimation and model validity tests.

In the next section a brief summary of all algorithms used in NLI (recursive least squares, extented recursive least squares, instrumental variable and suboptimal least squares) are discussed.

3.4.1) Least squares parameter estimation algorithm for the NARMAX models

Consider the NARMAX model represented by equation 3.6 and introduce the time delay d

$$y(t) = F^l_*[y(t-1), \ldots, y(t-n_y); u(t-d), \ldots, u(t-d-n_u+1)] \qquad (3.13)$$

where

$F_*[.]$ is some nonlinear function of u and y

n_y is the order of the lagged outputs

n_u is the order of the lagged inputs

l is the degree of nonlinearity

If $F_*[.]$ is assumed to be a polynomial function and defining V_1 = y(t-1) , V_2 = y(t-2),...,V_{ny} =y(t-n_y), V_{ny+1} =u(t-d),...,V_s =u(t - d - n_u + 1)

where

s= $n_y + n_u$.

Equation 3.13 can be represented as

$$y(t) = F_*^l[V_1, V_2, \ldots, V_{s-1}, V_s] \tag{3.14}$$

Expanding equation 3.14

$$y(t) = \sum_{i=1}^{s} \theta_i V_i + \sum_{i=1}^{s}\sum_{j=1}^{s} \theta_{ij} V_i V_j + \ldots\ldots$$
$$\ldots\ldots + \underbrace{\sum_{i=1}^{s}\sum_{j=1}^{s} \cdots \sum_{m=1}^{s}\sum_{n=1}^{s} \theta_{ij\ldots mn} V_i V_j \cdots V_m V_n}_{l \text{ times}} \tag{3.15}$$

If the output y(t) is assumed to be corrupted by zero mean additive noise e(t), the output z(t) may be written as

$$z(t) = y(t) + e(t) \tag{3.16}$$

Substituting equation 3.16 into 3.15

$$\begin{aligned} z(t) = {} & \sum_{i=1}^{n_y} \theta_i (z(t-i) - e(t-i)) + \sum_{i=1}^{n_u} \theta_{ny} (u(t-d-i+1)) \\ & + \sum_{i=1}^{n_y}\sum_{j=1}^{n_y} \theta_{ij} (z(t-i) - e(t-i))(z(t-j) - e(t-j)) \\ & + 2\sum_{i=1}^{n_y}\sum_{j=1}^{n_u} \theta_{i,ny+j} (z(t-i) - e(t-i))(u(t-d-j+1) \\ & + \sum_{i=1}^{n_u}\sum_{j=1}^{n_u} \theta_{ny+i,ny+j} (u(t-d-i+1)(u(t-d-j+1)) + \ldots\ldots \\ & \ldots\ldots \textit{higher order terms up to the degree } l + e(t) \end{aligned} \tag{3.17}$$

Equation 3.17 can be represented as

$$\begin{aligned} z(t) = {} & F_*^l[z(t-1),\ldots,z(t-n_y); u(t-d), \ldots, u(t-d-n_u+1), \\ & ; e(t-1), \ldots, e(t-n_y)] + e(t) \end{aligned}$$

In general, noise is nonlinear and can not be added to the output so equation 3.17 becomes

$$\begin{aligned} z(t) = {} & F_*^l[z(t-1),\ldots,z(t-n_y); u(t-d),\ldots,u(t-d-n_y+1); \\ & e(t-1),\ldots,e(t-n_y),e(t)] \end{aligned} \tag{3.18}$$

In general, equation 3.18 can be used to represent a wide class of nonlinear systems whatever the noise. When the system is nonlinear, it is impossible to solve equation 3.18. Re-arranging equation 3.18 into a prediction error model gives the

general form.

$$z(t) = F^{l}_{*}[z(t-1), \ldots, z(t-n_y); u(t-d), \ldots, u(t-d-n_u+1), \\ ; \varepsilon(t-1), \ldots, \varepsilon(t-n_y)] + \varepsilon(t) \tag{3.19}$$

where

$\varepsilon(t)$ is the prediction error equal to z(t)-$\hat{z}(t)$

and

$$E\left[\frac{\varepsilon(t)}{z(t-1), z(t-2), \ldots, u(t-d), \ldots}\right] = 0$$

Expanding equation 3.19

$$z(t) = G^{zu}[z(t-1),\ldots,z(t-n_y),\ldots;u(t-d),\ldots,u(t-d-n_u+1)] \\ + G^{zu\varepsilon}[z(t-1),\ldots,z(t-n_y);u(t-d),\ldots,u(t-d-n_y+1);\varepsilon(t-1),\ldots,\varepsilon(t-n_y)] \\ + G^{\varepsilon}[\varepsilon(t-1),\ldots,\varepsilon(t-n_y) + \varepsilon(t)] \tag{3.20}$$

where

G^{zu} is a polynomial function of z and u only.

$G^{zu\varepsilon}$ is a polynomial function of all cross product terms involving $\varepsilon(t)$

and

G^{ε} is a polynomial function of prediction errors only.

Separating out the unknown parameters of equation 3.20

$$z(t) = \Psi^{T}(t)\Theta(t-1) + \varepsilon(t) = [\Psi^{T}_{zu}(t)\Psi^{T}(t)\Psi^{T}_{\varepsilon}(t)]\begin{bmatrix} \Theta_{zu}(t-1) \\ \Theta_{zu\varepsilon}(t-1) \\ \Theta_{\varepsilon}(t-1) \end{bmatrix} \tag{3.21}$$

where

$$\left.\begin{aligned} G^{zu}[.] &= \Psi^{T}_{zu}(t)\Theta_{zu}(t-1) \\ G^{zu\varepsilon}[.] &= \Psi^{T}_{zu\varepsilon}(t)\Theta_{zu\varepsilon}(t-1) \\ G^{\varepsilon}[.] &= \Psi^{T}_{\varepsilon}(t)\Theta_{\varepsilon}(t-1) \end{aligned}\right\} \tag{3.22}$$

grouping all terms involving $\varepsilon(t)$ and defining

$$\xi(t) = \Psi^{T}_{zu\varepsilon}(t)\Theta_{zu\varepsilon}(t-1) + \Psi^{T}_{\varepsilon}(t)\Theta_{\varepsilon}(t-1) + \varepsilon(t) \tag{3.23}$$

gives

$$z(t) = \Psi^{T}_{zu}(t)\Theta_{zu}(t-1) + \xi(t) \tag{3.24}$$

equation 3.23 above shows that $\xi(t)$ is highly correlated with the elements of $\psi_{zu}^T(t)$. Therefore application of least squares will lead to biased estimates.

3.4.2) Recursive least squares

From equation 3.23 minimising the cost function J,

where

$J = \sum_{t=1}^{N} \xi^2(t)$ leads to the well known least squares solution

$$\hat{\Theta}_{zu} = [\Psi_{zu}^T \Psi_{zu}]^{-1} \Psi^T \underline{z} \tag{3.25}$$

where

$$\Psi_{zu} = [\psi_{zu}^T(1), \psi_{zu}^T(2), \ldots, \psi_{zu}^T(N)] \tag{3.26}$$

and

$$\underline{z} = [z(1), z(2), \ldots, z(N)]^T \tag{3.27}$$

equation 3.30 can be transformed into a recursive estimator which is simple and can be implemented using the unified representation algorithm by Sodestrom et al, (1974)

$$\hat{\Theta}(t+1) = \hat{\Theta}(t) + K(t+1)\varepsilon(t+1) \tag{3.28}$$

$$K(t+1) = \frac{P(t)\Phi(t+1)}{\Lambda(t+1) + \psi^T(t+1)P(t)\Phi(t+1)} \tag{3.29}$$

$$P(t+1) = \frac{\left[P(t) - \dfrac{P(t)\Phi(t+1)\psi^T(t+1)P(t)}{\Lambda(t+1) - \psi^T(t+1)P(t)\Phi(t+1)}\right]}{\Lambda(t+1)} \tag{3.30}$$

$$\Lambda(t+1) = \Lambda_0\Lambda(t) + (1-\Lambda_0) \tag{3.31}$$

$$\xi(t) = z(t+1) - \psi^T(t+1)\hat{\Theta}(t) \tag{3.32}$$

where

$\Lambda(t)$ is the forgetting factor

and

$\hat{\Theta}(t)$, $\xi(t)$, $\psi(t)$ and $\Phi(t)$ are given below with the following definitions

$$\left.\begin{aligned} \Theta(t) &= \Theta_{zu}(t) \\ \psi^T(t+1) &= \psi^T_{zu}(t+1) \\ \Phi^T(t+1) &= \psi^T(t+1) \end{aligned}\right\} \tag{3.33}$$

This algorithm computes the initial estimates of the unknown parameters and the error vector of the NARMAX model.

3.4.3) Extended recursive least squares (ERLS)

Panuska's method or extended least squares algorithm for linear systems has been widely used (Goodwin et al, 1977).

The ERLS can be implemented for the NARMAX model using equation 3.28 to 3.32 with

$$\left.\begin{aligned} \Theta(t) &= [\Theta_{zu}(t)\Theta_{zu\varepsilon}(t)\Theta_{\varepsilon}(t)] \\ \psi(t+1) &= [\psi^T_{zu}(t+1)\psi^T_{zu\varepsilon}(t+1)\psi^T_{\varepsilon}(t+1)] \\ \Phi(t+1) &= [\psi(t-1)] \end{aligned}\right\} \tag{3.34}$$

From equation 3.34 the noise or prediction error parameters are included in $\Theta(t)$, this to eliminate bias in the estimates.

If the nonlinearities within the system are severe, the dimension of $\hat{\Theta}(t)$ increases rapidly. To limit the dimension of the vector $\hat{\Theta}(t)$ both instrumental algorithms and sub-optimal least squares algorithms were developed (Goodwin et al, 1977), (Young, 1970), (Billings et Voon, 1984).

3.4.4) Instrumental variables

The principle of instrumental variables has been applied to linear systems identification, but it has also been extended to the NARMAX models.

The off line description of these algorithms is based on the linear model

$$Y = \Phi\Theta + e \tag{3.35}$$

and the selection of an instrument matrix V^T such that

$$\left.\begin{aligned} \lim \frac{1}{N} V^T \Phi &= R \\ \lim \frac{1}{N} V^T (Y - \Phi\Theta_0) &= 0 \end{aligned}\right\} \tag{3.36}$$

where

Θ_0 denotes the true parameter vector

lim refers to limit in probability

N sequence of N output measurements

The above conditions ensure that the estimate

$$\hat{\Theta} = (V^T\Theta)^{-1} V^T Y \tag{3.37}$$

is unbiased.

There are a number of ways of selecting the matrix V^T so that equation 3.36 is satisfied. The most popular is either delayed inputs to form V^T or define V^T to have the same structure as Φ^T but with the measured outputs replaced by the predicted outputs. This latter algorithm is often referred to as the auxiliary model algorithm. Unfortunately instrumental variables can only be applied to nonlinear systems when certain properties of the system noise are satisfied.

Consider the NARMAX model of equation 3.24

$$z(t) = \Psi_{zu}^T(t)\Theta_{zu}(t-1) + \xi(t) \tag{3.38}$$

where

$$\xi(t) = \Psi_{zu\varepsilon}^T(t)\Theta_{zu\varepsilon}(t-1) + \Psi_{\varepsilon}^T(t)\Theta_{\varepsilon}(t-1) + \varepsilon(t) \tag{3.39}$$

For a sequence of N output measurements define

$$\left.\begin{aligned} \underline{z} &= [z(1), \ldots, z(N)]^T \\ \underline{\Psi}_{zu} &= [\Psi_{zu}^T(1), \ldots, \Psi_{zu}^T(N)]^T \\ \underline{\xi} &= [\xi(1), \ldots, \xi(N)]^T \end{aligned}\right\} \tag{3.40}$$

to yield the description

$$\underline{z} = \underline{\Psi}_{zu}\Theta_{zu} + \underline{\xi} \tag{3.41}$$

Forming an instrumental matrix V^T using the auxiliary model algorithm gives the instrumental variable estimate

$$\hat{\Theta}_{zu} = (V^T\Psi_{zu})^{-1}V^T\underline{z} \tag{3.42}$$

which will be biased whenever the process under test is nonlinear. A similar conclusion follows even if V^T is formed from delayed inputs. This arises because the terms of the form

$$\lim\frac{1}{N}V^T[z^i u^j \varepsilon^k] \tag{3.43}$$

for some i, j, k will not in general vanish even when $\varepsilon(k)$ is a zero mean white noise sequence (Billings and Voon, 1984). Equation 3.43 will however always tend to zero for the special case of the NARMAX model. When the noise terms are represented within the model as a linear map L[.]

$$\begin{aligned} z(t) = F^l[&z(t-1), \ldots, z(t-n_y); u(t-d), \ldots, u(t-d-n_u+1)] \\ &+ L[\varepsilon(t-1), \ldots, \varepsilon(t-n_\varepsilon)] + \varepsilon(t) \end{aligned} \tag{3.44}$$

These conditions can always be tested using the model validity tests described in section 3.3.2.

The off-line instrumental variable estimator as defined in equation 3.41 can be implemented using the recursive equations 3.28 to 3.32 with the following definitions of the matrices:

$$\left.\begin{aligned} \Psi^T(t+1) &= [\Psi^T_{zu}(t+1)] \\ \hat{\Theta}(t) &= [\Theta_{zu}(t)] \\ zz(t) &= [\Psi_{\hat{z}u}(t+1)] \end{aligned}\right\} \tag{3.45}$$

3.4.5) Suboptimal least squares

If the noise model is included in the estimator, the number of parameters to be estimated in the NARMAX model increases significantly. It would therefore be advantageous if unbiased process parameter estimates could be obtained without specifically estimating a noise model. This can be achieved for the NARMAX model if the noise is considered to be an additive signal at the output by using a new suboptimal least squares algorithm

Consider the NARMAX model given by equation 3.7

$$z(t) = F^l_*[z(t-1),...,z(t-n_y);u(t-d),...,u(t-d-n_u+1);$$
$$\varepsilon(t-1),...,\varepsilon(t-n_y)] + \varepsilon(t) \tag{3.46}$$

Expansion of equations 3.46 leads to a considerable number of cross product terms between the noise and the measured input and the output signals. If however y(t) is noise-free, then y(t) could be monitored and equation 3.45 could be represented as

$$z(t) = F^l[y(t-1), \ldots, y(t-n_y);u(t-d),...,u(t-d-n_u+1)] + e(t) \tag{3.47}$$

and all the cross-product terms are eliminated. Although the signal y(t) is unavailable for measurement it can be estimated by recursively computing the predicted output using

$$\hat{y}(t) = \Psi^T_{\hat{y}u}(t)\hat{\Theta}_{yu}(t-1) \tag{3.48}$$

The noise-free output y(t) in equation 3.47 is therefore replaced by the estimate $\hat{y}(t)$.

The algorithm is simple to implement using equation 3.28 to 3.32 with the following definitions

$$\left.\begin{aligned} \Psi^T(t+1) &= \Psi^T_{\hat{y}u}(t+1) \\ \Theta(t) &= \Theta_{zu}(t) \\ zz(t) &= \Psi(t) \end{aligned}\right\} \tag{3.49}$$

It is important to remember that this algorithm is applicable only when the noise is additive at the output. Otherwise, the cross-product noise terms which appear in $\Psi^T_{zu\varepsilon}(t)$ may induce bias. This situation can be detected using the model validity tests.

The noise in the nonlinear systems whether internal or output additive complicates the estimation problem. So a modified extended least squares and a new suboptimal least squares algorithms were introduced as one possible solution to yield unbiased estimates in the presence of multiplicative noise terms. Instrumental variables which are widely used for linear systems yield biased estimates whenever the noise model cannot be expressed as a linear map.

If the extended least squares algorithm is used, the conditions of equation 3.22 must be satisfied. Instrumental variables and suboptimal least squares do not specifically estimate a noise model and consequently the residuals may be coloured.

From all of the five algorithms the extended recursive least squares is perhaps the best. This is probably due to the inclusion of the additive noise in the specification of the NARMAX model. But in some cases, due to the limitation of the NLI package, the extended recursive least squares can not be used. In these situation, instrumental variable or suboptimal recursive least squares has been used. Since both algorithms do not include the coefficients of the noise model in the parameter vector, the number of terms in the model is very much less than that in the extended recursive least squares case.

3.5) Use of the NLI package

The NLI program is written in FORTRAN 77. It has been implemented on the PERKIN-ELMER 3220 computer in the Department of Control Engineering, for multi-user access via any graphical terminal. Each option is selected by typing three characters. The data are supplied to the program. Then a nonlinear identification algorithm is selected. At the conclusion, the user may store the calculated data at the end of a run (predicted output, deterministic prediction errors and residuals) in the new data file specified by the user. This new data file can be used to interact with SPAID (section 3.6.3). If desired, a display of plots of input and output data may be chosen. After the data are read by the program, and before nonlinear identification is applied, a structure detection test is used to determine if the process under test exhibits nonlinear characteristics which warrant a nonlinear model (section 3.3.1).

The next step is to select one of the four ' least squares ' type recursive algorithms contained in the package, these are : Recursive least squares, Auxiliary

instrumental variable, Extended Recursive least squares and Suboptimal Recursive least squares.

Once one of these algorithms is selected, the initial specifications for the NAR-MAX model are entered which are : the order of the output, the order of the input, order of the prediction error, the time delay and the degree of nonlinearity of input and output. The program will display the coefficients which define the NARMAX model. Terms can be selected manually. The use of the forward and backward regression gives the terms which are significant. Also extra terms can be forced onto the model to improve the MVT.

If the Fisher-Ratio is used, the forward and backward stepwise regressions are often selected together because any insignificant terms that are included in the model at the early stages may become insignificant when other terms are included at a later stage. These insignificant terms are deleted so that only the significant terms are retained in the final model. In both forward and backward regressions, the critical value of the Fisher-Ratio is assumed to be 4.0.

Finally, a display of five options is given by the program namely (print coefficients of the model, print covariance, plot the residuals (input, output, predicted output, residuals and deterministic errors) and MVT option based on correlation functions).

Each algorithm requires two channels of data : input and output channels. The two channels must be stored as two columns in indexed 80 bytes/record data file.

The preliminary analysis of the data file involved estimating the coefficients in a linear model first, of varying the initial specification, and with various time delays, and analysing the loss function for each case.

3.5.1) Example of use of the NLI package

In the present section, the use of the NLI package will be illustrated using an example of a heat exchanger model (Billings and Fadzil, 1984).

The heat exchanger consists of a radiator through which heated water is passed and a fan which blows air across the radiator. Water is pumped through the radiator around a closed loop which includes a heater tank (Figure 3.4). The aim is to control the temperature drop across the radiator together with the air flow rate across it by adjusting the input to the heater and the fan. A block diagram of the system is shown in Figure 3.5. With reference to Figure 3.5 extensive experimentation has shown that G_{22} and G_{12} are linear while G_{11} is nonlinear. Therefore only the nonlinear loop will be considered.

3.5.2) System identification

Having entered the data into the NLI the structure detection tests showed that $\Phi^2_{zz}(\tau)$ is well outside the confidence bounds (Figure 3.6), confirming that the system is nonlinear (loop G_{11}).

Using a prediction error estimation algorithm coupled with a stepwise regression procedure the model coefficients were estimated. The initial specification for the NARMAX model was :

Order of the lagged output,	$n_y = 2$
Order of lagged input,	$n_u = 2$
Order of lagged prediction error,	$n_e = 2$
Delay of the input,	d = 1
Degree of nonlinearity of input and output,	l = 3
Degree of nonlinearity of prediction error,	$l_e = 2$

With this initial specification the number of terms in the model was 83. Using the Fisher-Ratio test with 95% confidence bounds, the following model was obtained :

$$\begin{aligned} z(t) = {} & 2.072 + 0.9158z(t-1) + 0.4788u(t-1) - 0.01572z^2(t-1) \\ & - 0.0113u^2(t-1) - 0.00244z^2(t-1)u(t-1) \\ & - 0.002239u^3(t-1) + \varepsilon(t) \end{aligned} \tag{3.50}$$

The estimated model (equation 3.12) was good but the MVT, shown in Figure 3.7, indicates that the model is deficient in some way. $\Phi_{u'^2\xi}(\tau)$ is inside the 95%

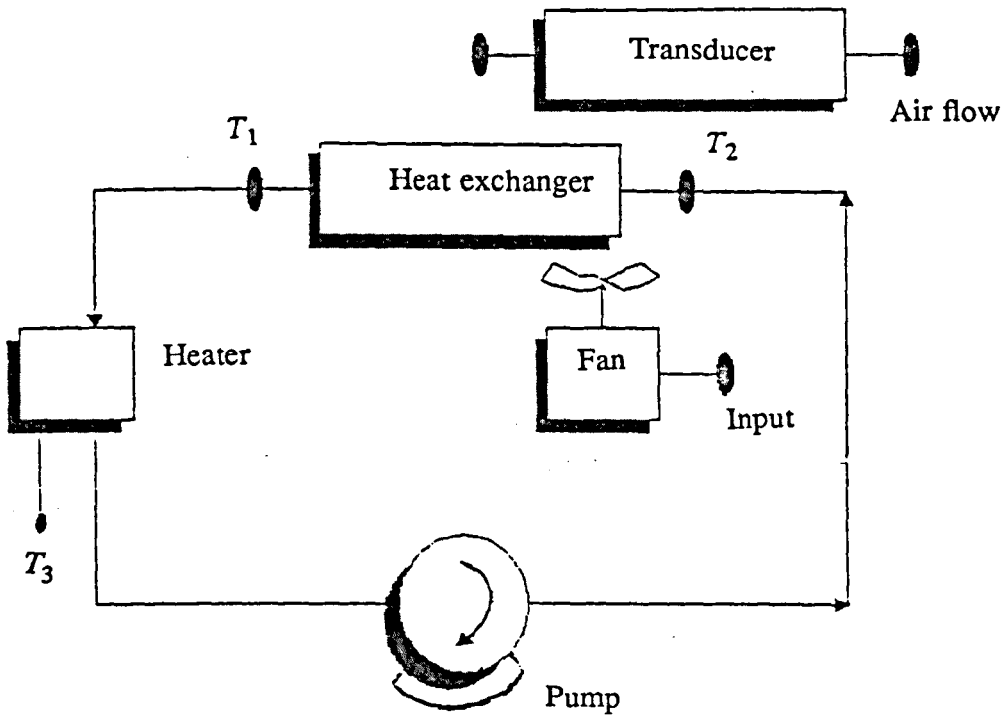


Figure: 3.4

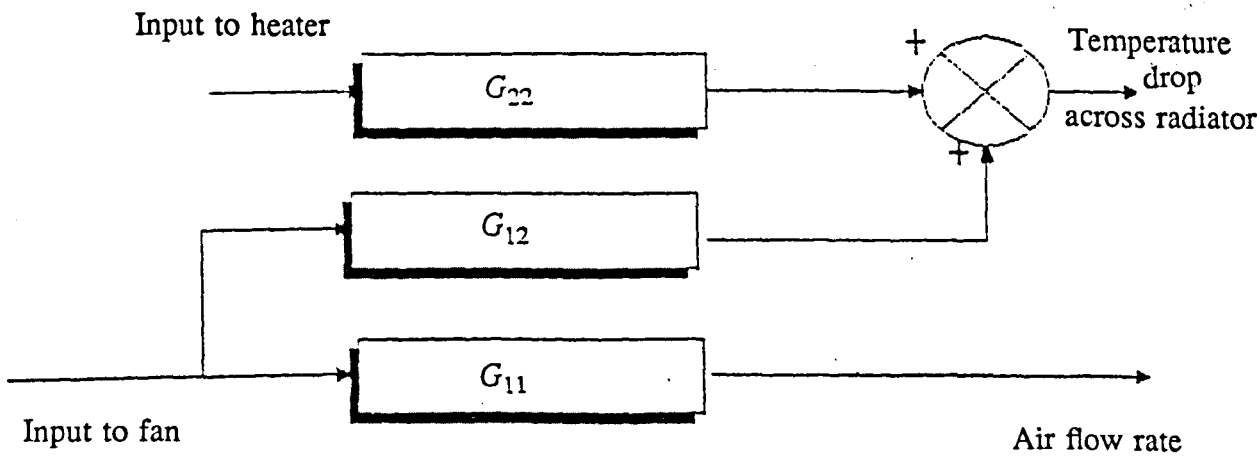


Figure: 3.5

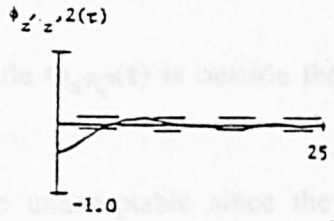


Figure 3.6: Structure detection test

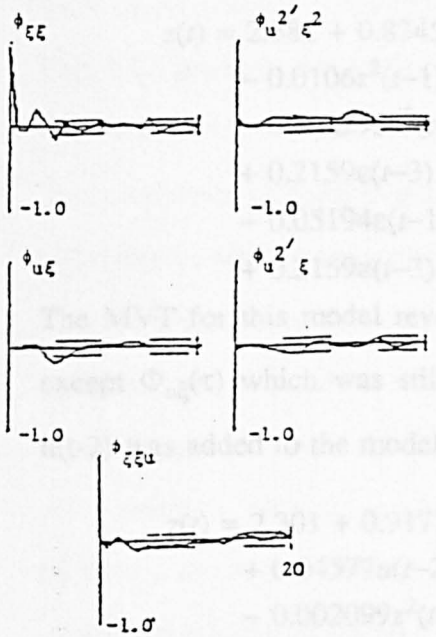


Figure 3.7: Model validity test

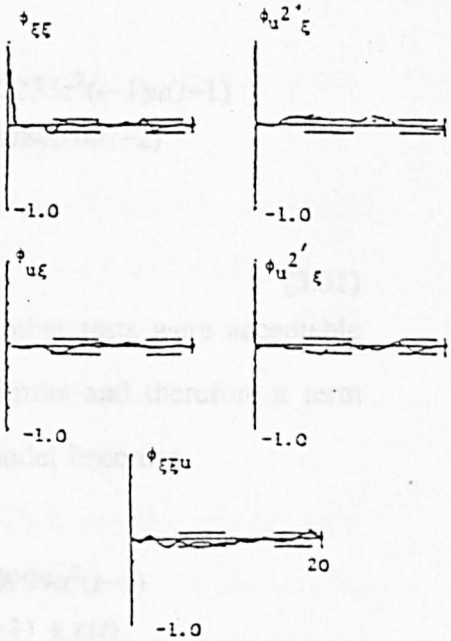


Figure 3.8: Model validity test

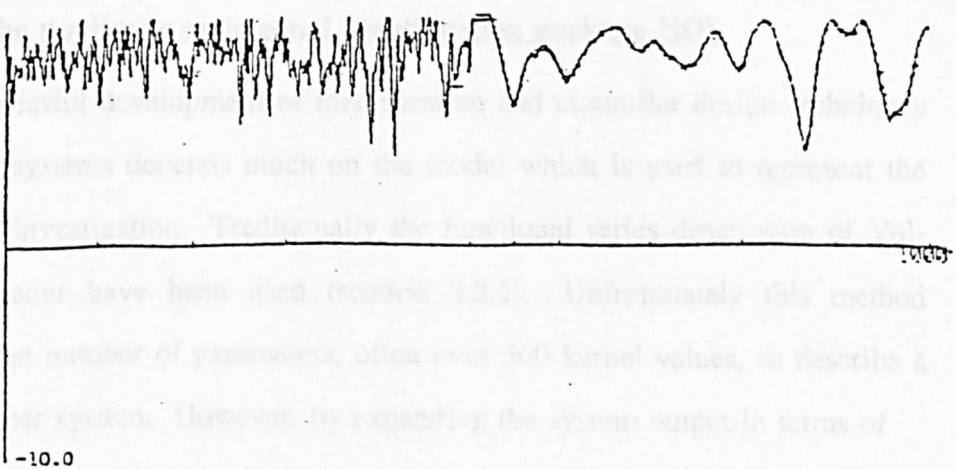


Figure 3.9: Process and predicted output superimposed

confidence bounds while $\Phi_{u'^2\zeta^2}(\tau)$ is outside the confidence bounds for lag 11 and lag 12.

$\Phi_{\zeta\zeta}(\tau)$ and $\Phi_{u\zeta}(\tau)$ are unacceptable since they are outside the 95% confidence bounds, this means that addition of higher order noise model is required to improve the model and therefore equation 3.50 becomes

$$\begin{aligned} z(t) = {} & 2.381 + 0.8345z(t-1) + 0.4828u(t-1) \\ & - 0.0106z^2(t-1) - 0.00867u^2(t-1) - 0.002235z^2(t-1)u(t-1) \\ & - 0.002293u^3(t-1) - 0.05194\varepsilon(t-1) - 0.008461\varepsilon(t-2) \\ & + 0.2159\varepsilon(t-3) + 0.08243\varepsilon(t-4) + \varepsilon(t) \\ & - 0.05194\varepsilon(t-1) - 0.008461\varepsilon(t-2) \\ & + 0.2159\varepsilon(t-3) + 0.08243\varepsilon(t-4) \end{aligned} \tag{3.51}$$

The MVT for this model revealed that $\Phi_{\xi\xi}(\tau)$ and all other tests were acceptable except $\Phi_{u\xi}(\tau)$ which was still outside the confidence limits and therefore a term u(t-2) was added to the model of equation 3.49. The model becomes

$$\begin{aligned} z(t) = {} & 2.301 + 0.9173z(t-1) + 0.449u(t-1) \\ & + 0.04577u(t-2) - 0.01889z^2(t-1) - 0.00999u^2(t-1) \\ & - 0.002099z^2(t-1)u(t-1) - 0.002434u^3(t-1) + \varepsilon(t) \\ & - 0.004\varepsilon(t-1) + 0.0380\varepsilon(t-2) + 0.2745\varepsilon(t-3) \\ & + 0.1037\varepsilon(t-4) \end{aligned} \tag{3.52}$$

The MVT is shown in Figure 3.8. The correlation function are well within the confidence bounds. The predicted output of the final model of equation 3.52 is shown in Figure 3.9.

3.6) Use of the nonlinear orthogonal identification package NOI

The successful development of identification and controller design techniques for nonlinear systems depends much on the model which is used to represent the system under investigation. Traditionally the functional series description of Volterra and Wiener have been used (section 3.2.1). Unfortunately this method requires a large number of parameters, often over 500 kernel values, to describe a simple nonlinear system. However, by expanding the system output in terms of

past inputs and outputs using a NARMAX model a very concise representation for a wide class of nonlinear systems can be obtained to ease the problems associated with the functional series methods.

A new orthogonal parameter estimation algorithm (appendix B) was derived by (Billings et al, 1987) and implemented by (Billings and Tsang, 1987) on the SUN-mycrosystems at Sheffield University.

This NOI package consists of several suites of programs for data generation, structure detection, parameter estimation and model validity tests.

The parameter estimation program is composed of two algorithms : the extended orthogonal estimation algorithm and the prediction error algorithm. The first algorithm provides a quick way of fitting a model, while the second is slow but gives a much better estimate. Therefore, if an optimised model is required, one can pass the orthogoal estimate into the prediction error algorithm.

The program displays a menu of options. Each algorithm requires two channels of data : input and output. The maximum number of data pairs is 1024. The maximum number of terms in the final model is 100, the maximum degree of nonlinearity is 5, the maximum lags $n_y + n_u$ and n_e must both be equal to 100 and there is no maximum for time delay.

where

n_y represents the number of lags in the output

n_u represents the number of lags in the input

n_e represents the number of lags in the prediction error

3.7) Other packages

This section gives a brief summary of other packages used during the course of this research. These packages are PSI, PSICON and SPAID.

3.7.1) PSI

This is an interactive simulation language implemented in FORTRAN on the

large multi-user interactive PERKIN-ELMER 3240 in the Department of Control Engineering. This package has been developed by Van Den Bosh (1979) at Delft University of Technology, the Netherlands.

It is a block-oriented simulation language. The package contains over 40 different block types including standard continuous dynamic elements such as a integrators, discrete elements such as zero-order-hold, time delay, a three-term (PID) controller, nonlinear elements such as dead-zone and saturation devices. Also available in PSI is a hill climbing optimization procedure used to optimize up to nine parameters which is very important for the design of optimal (PID) controllers.

If graphical output is required then up to four variables may be printed on the screen or the line printer. Each block has a maximum of three inputs and three parameters.

3.7.2) PSICON

The simulation program PSICON (Linkens et al, 1982) is an extension of the block-oriented simulation program PSI. The parent language PSI is extended to make use of multitasking, to allow for inclusion of FORTRAN subroutines to give increased flexibility for incorporation of modern control algorithms. This interaction is achieved via message transfer between the two tasks using the multiuser operating system of the PERKIN-ELMER 3240 on which the program runs.

PSI runs on a Tektronix 4010 grphics terminal screen which provides the continuous plotting of up to four variables and the external FORTRAN segment and overall control of the simulation is managed from an other terminal (Menad, 1984).

3.7.3) SPAID

The SPAID package consists of several suites of programs for data handling and modification, parameter estimation, model validity tests, simulation and plotting.

The package has been design by Billings and Sterling, (1979). It is an interactive package written in FORTRAN 77, for the identification of linear single-input single-output time invariant systems. It has been implemented on the PERKIN-ELMER 3240 mini computer for multiuser access via any graphics terminal (Billings et al, 1979) and (Batey et al, 1975).

CHAPTER 4

Nonlinear Identification of Vecuronium and Atracurium Dynamics

4.1) Introduction

An input/output model of the process dynamics for muscle relaxant drugs comprises two parts. One is the linear pharmacokinetics (distribution of drug into the blood) and which can be described by a number of time constants, the other is the so called pharmacodynamics which describe the effect mechanism of the drug, and which can be represented by a static nonlinear characteristic.

The emphasis in this work is on obtaining an input/output dynamic model which can be used for the design of controller algorithms for on-line infusion of drugs during surgery.

The present work is the sequel to previous studies which used PRBS excitation under open loop conditions to obtain a linearized model for the drug Pancuronium administered in animal trials (Linkens et al, 1982).

A two-stage identification process was used to estimate the nonlinear pharmacodynamics for Vecuronium employing the linearized model obtained using PRBS excitation mentioned above (Linkens and Asbury, 1985).

Identification of nonlinear difference equation models for the Vecuronium and Atracurium drugs is considered based on the NARMAX model representation providing a general description of the process, but giving no indication as to the underlying physiological structure. It can however, be used with short lengths of data and under a variety of signal conditions. For these reasons, this technique is first applied to the experimental data under closed-loop control using Vecuronium and Atracurium drug infusion and the evoked EMG measurements.

This chapter presents two main studies of neuromuscular blockage using both Vecuronium and Atracurium.

The first study is to fit a nonlinear model to the data of Vecuronium data using the NLI package with an extended recursive least squares algorithm documented in the previous chapter.

After the initial direct identification of the data, initial conditions were introduced to account for the effect of the bolus injection administrated at the begining of the trial.

In the second study, the NLI and NOI (Nonlinear Orthogonal Identification) packages are used to fit a nonlinear model to the Atracurium data.

4.2) Data acquisition and conditioning for Vecuronium

The experiments were performed on healthy mongrel dogs (Linkens and Asbury, 1985). Anaesthetic drugs were injected intravenously into the left leg with methohexitone 5mg/ml, and maintained using a constant infusion of doperidol 0.4mg/Kg/hr and fentanyl 6 μ g/Kg/hr. Automatic ventilation with O_2/N_2O was used throughout the experiments, and the tidal volume adjusted to maintain an end-exhaled CO_2 concentration of 4.2%. The level of relaxation was obtained from a rectified, integrated EMG (RIEMG) signal evoked using a 0.1Hz stimulus on the front leg and the muscle relaxant drug (Vecuronium) was continuously infused into the right leg via a roller pump driven by a d.c motor. The experimental set up is illustrated in Figure 4.1. The data obtained from this experiment are shown in Figure 4.2.

Prior to the identification studies the data were normalised, so that an RIEMG value of 1 represented 100% relaxation (Figure 4.3). The data were smoothed by a two points averaging procedure, giving 122 data points, as shown in Figure 4.4 with an equivalent sampling period of 20 seconds.

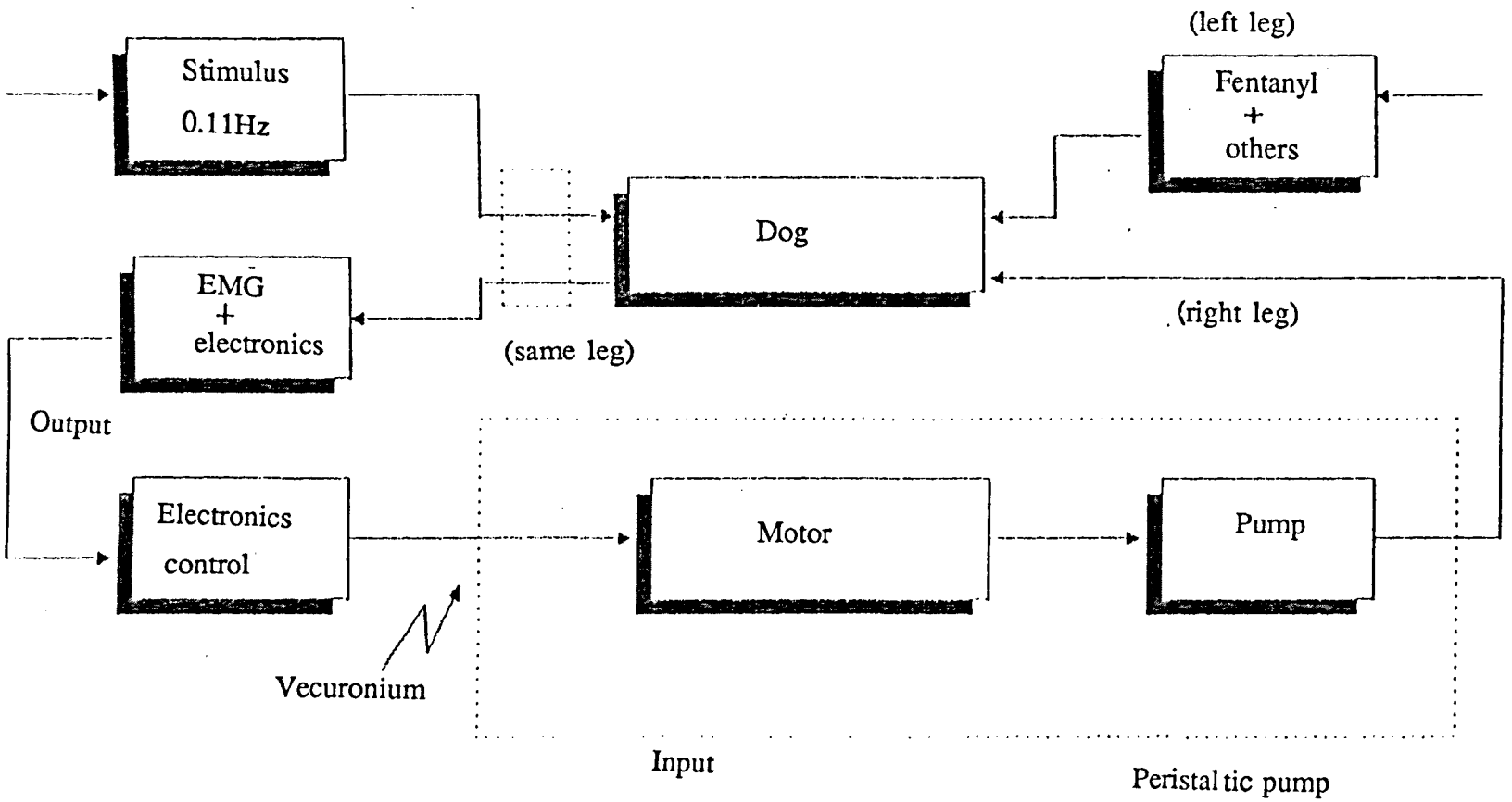


Figure 4.1: Data acquisition for Vecuronium

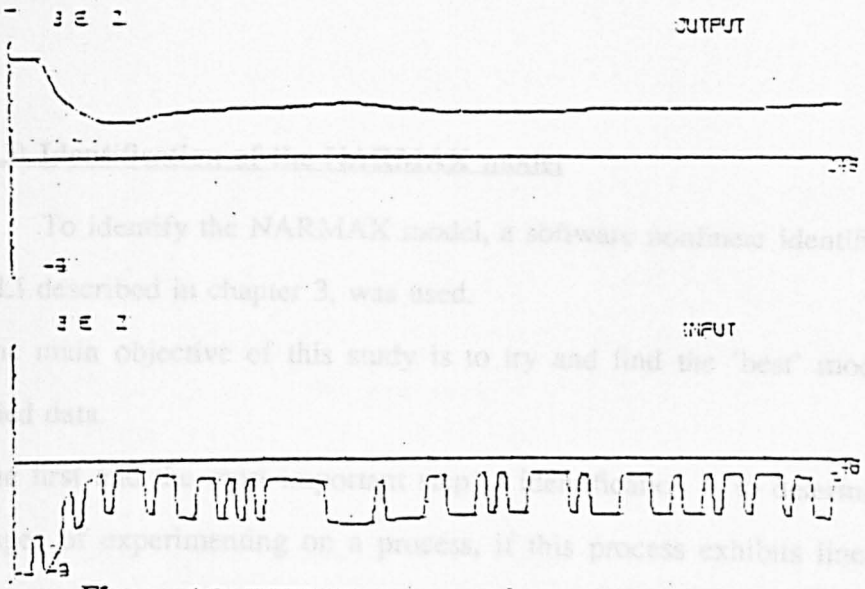


Figure 4.2: RIEMG recording from manually-controlled infusion

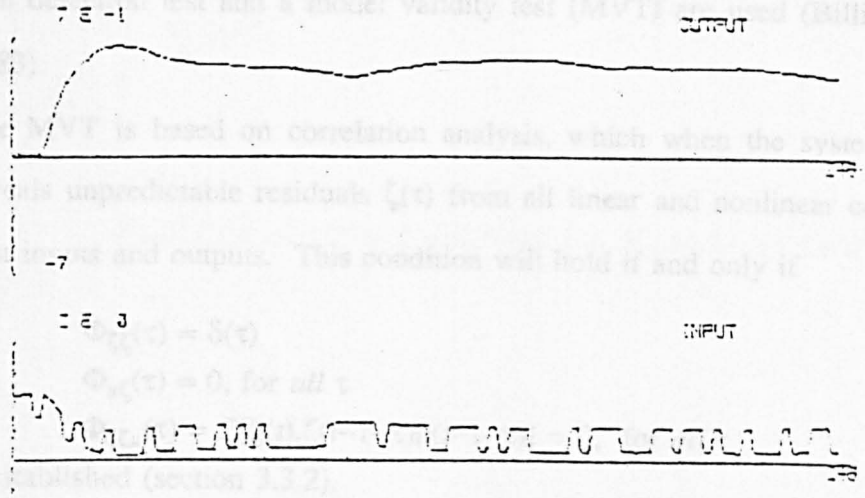


Figure 4.3: Representation of RIEMG and infusion of the actual data.

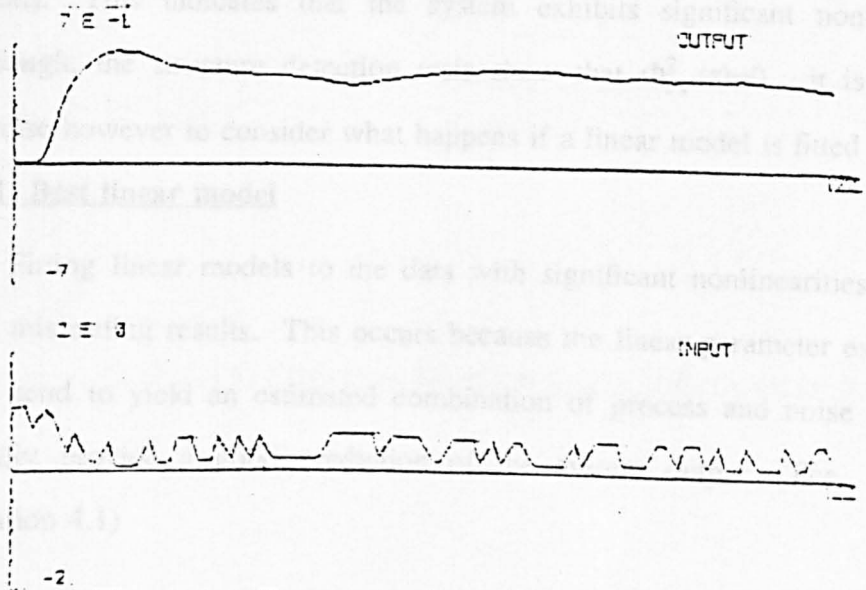


Figure 4.4: Representation of RIEMG and infusion of sensible data (after smoothing)

4.3) Identification of the NARMAX model

To identify the NARMAX model, a software nonlinear identification package, NLI described in chapter 3, was used.

The main objective of this study is to try and find the 'best' model for the supplied data.

The first and the most important step in identification is to determine in the early stages of experimenting on a process, if this process exhibits linear or nonlinear characteristics which warrant linear or nonlinear models. To achieve this a structure detection test and a model validity test (MVT) are used (Billings and Voon, 1983).

The MVT is based on correlation analysis, which when the system is nonlinear reveals unpredictable residuals $\zeta(\tau)$ from all linear and nonlinear combinations of past inputs and outputs. This condition will hold if and only if

$$\Phi_{\zeta\zeta}(\tau) = \delta(\tau)$$
$$\Phi_{u\zeta}(\tau) = 0, \text{ for } all\ \tau$$
$$\Phi_{\zeta\zeta u}(\tau) = E[\zeta(t),\zeta(t-1-\tau)u(t-1-\tau)] = 0, \text{ for } all\ \tau$$

is established (section 3.3.2).

The structure detection test of both data of Figures 4.3 and 4.4 shown in Figure 4.5 and Figure 4.6 respectively, show that $\Phi^2_{z'z'}(\tau)$ is well outside the confidence bounds. This indicates that the system exhibits significant nonlinear effects. Although, the structure detection tests show that $\Phi^2_{z'z'}(\tau)\neq 0$, it is a worthwile, exercise however to consider what happens if a linear model is fitted to the data.

4.3.1) Best linear model

Fitting linear models to the data with significant nonlinearities can produce very misleading results. This occurs because the linear parameter estimation routines tend to yield an estimated combination of process and noise model which visually provide a good prediction of the system output. The linear model (equation 4.1)

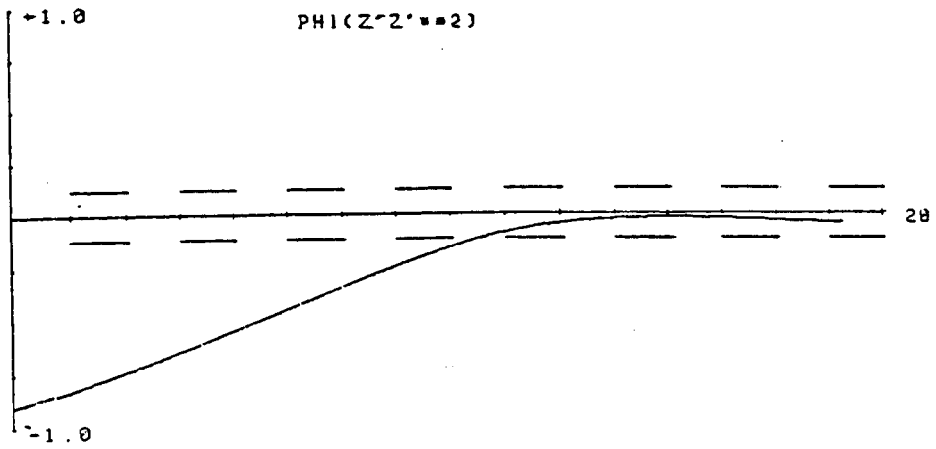


Figure 4.5: Structure detection test of the actual data

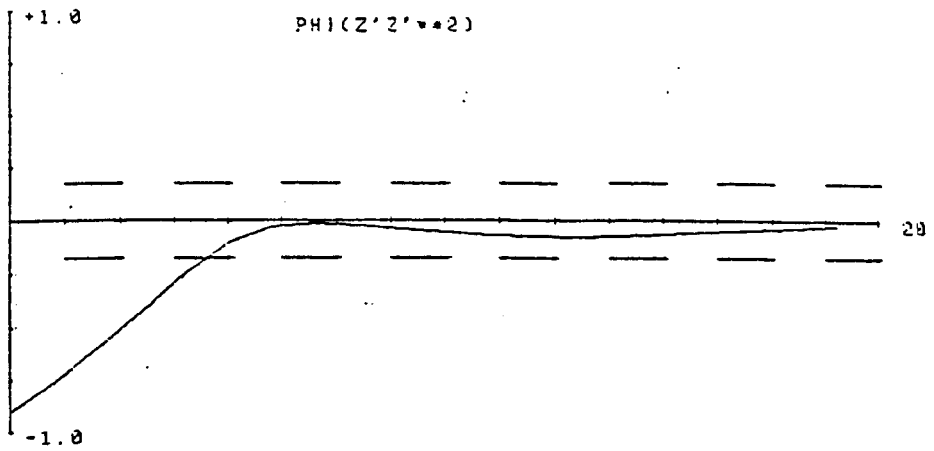


Figure 4.6: Structure detection test of the smoothed data

$$z(t) = 0.01484 + 1.354z(t-1) - 0.1952z(t-2) - 0.2306z(t-3) + 0.03431z(t-4) + 0.004898u(t-4) + 0.00322u(t-5) + 0.766210^{-4}u(t-6) + 0.005362u(t-7) - 0.04318e(t-1) \quad (4.1)$$

was estimated using an extended recursive least squares algorithm with the following specifications

The order of the lagged output,	$n_y = 4$
The order of the lagged input,	$n_u = 4$
The order of the noise,	$n_e = 1$
Delay of the input,	TD = 4
Degree of nonlinearity of the model,	l = 1

where

z and u are the system output and input at time t respectively and e is the noise. Extensive preliminary analysis of the data involved estimating the coefficients in linear models for different process and noise model orders and various time delays. The MVT for this linear model when the noise model order ($n_e = 1$) is shown in Figure 4.7. The plot of the residuals and the process and predicted outputs are illustrated in Figures 4.8 and 4.9 respectively.

From Figure 4.7, the cross-correlation analysis shows that $\Phi_{\zeta\zeta}(\tau)$ and $\Phi_{u\zeta}(\tau)$ are slightly outside the 95% confidence bounds. This indicates a deficiency in the noise model. To improve the MVT the noise model was gradually increased and the coefficients of the model were estimated using the ERLS with the following specification

The order of the lagged output,	$n_y = 4$
The order of the lagged input,	$n_u = 4$
The order of the noise,	$n_e = 4$
The delay of the input,	TD = 4
Degree of nonlinearity of the model,	l = 1

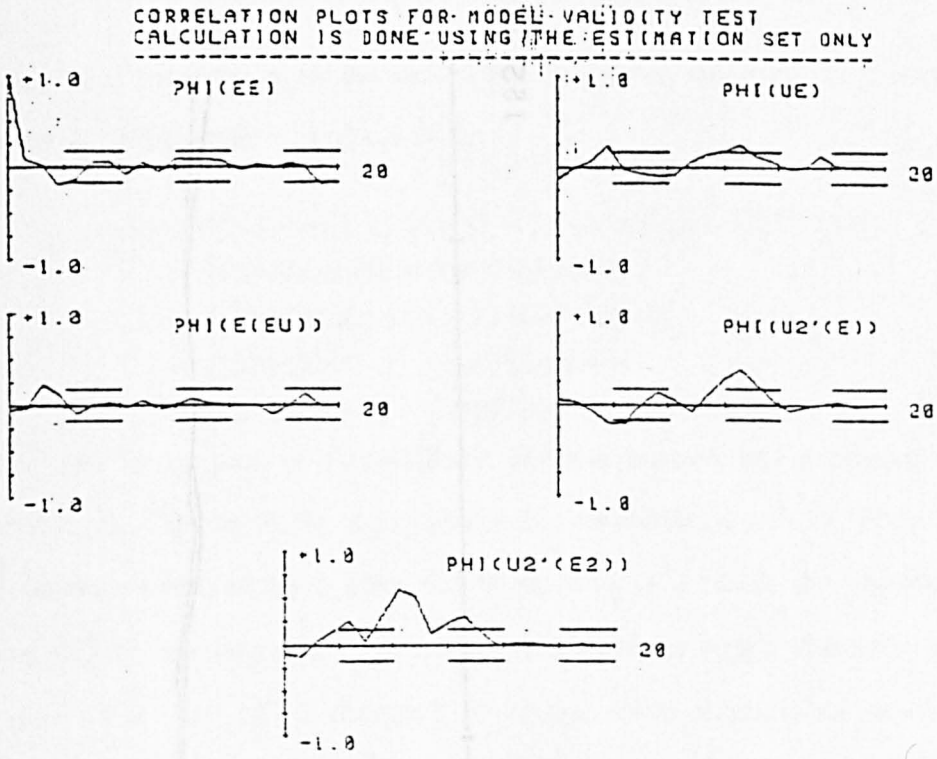


Figure 4.7: Model validity tests - best linear

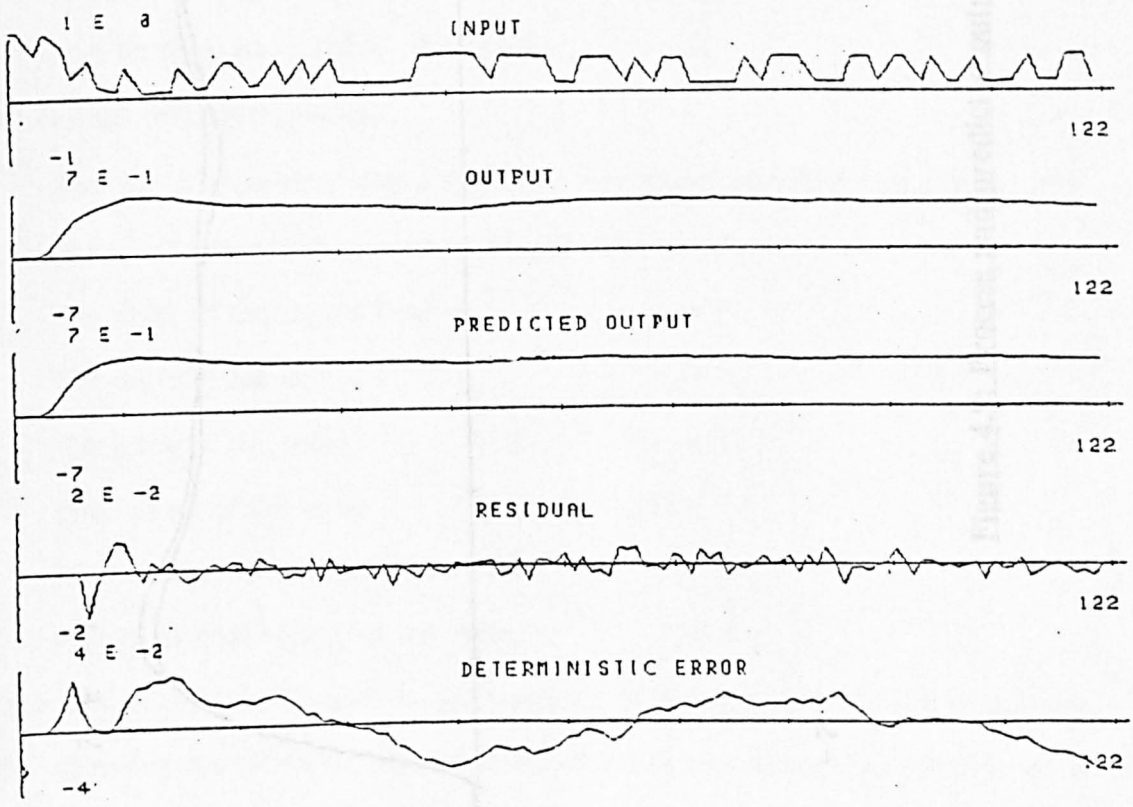


Figure 4.8: Plot of the residuals - best linear

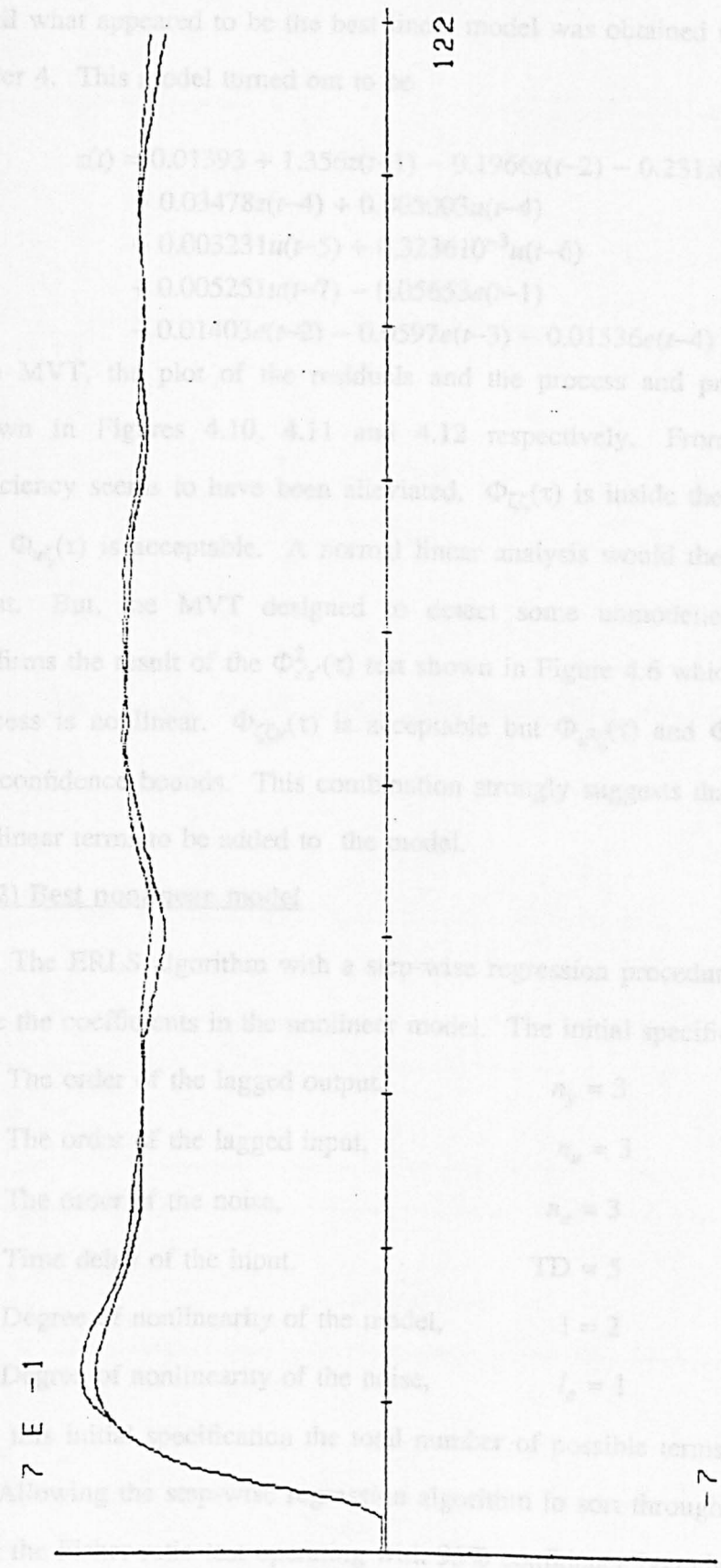


Figure 4.2: Process and predicted output superimposed - best linear

until what appeared to be the best linear model was obtained for a noise model of order 4. This model turned out to be

$$\begin{aligned} z(t) = {} & 0.01393 + 1.356z(t-1) - 0.1966z(t-2) - 0.231z(t-3) \\ & + 0.03478z(t-4) + 0.005003u(t-4) \\ & + 0.003231u(t-5) + 0.323610^{-3}u(t-6) \\ & + 0.005251u(t-7) - 0.05653e(t-1) \\ & - 0.01403e(t-2) - 0.0597e(t-3) - 0.01536e(t-4) \end{aligned} \tag{4.2}$$

The MVT, the plot of the residuals and the process and predicted outputs are shown in Figures 4.10, 4.11 and 4.12 respectively. From Figure 4.10, the deficiency seems to have been alleviated. $\Phi_{\zeta\zeta}(\tau)$ is inside the confidence bounds and $\Phi_{u\zeta}(\tau)$ is acceptable. A normal linear analysis would therefore finish at this point. But, the MVT designed to detect some unmodelled nonlinear terms, confirms the result of the $\Phi^2_{z'z'}(\tau)$ test shown in Figure 4.6 which indicates that the process is nonlinear. $\Phi_{\zeta\zeta u}(\tau)$ is acceptable but $\Phi_{u^2\zeta}(\tau)$ and $\Phi_{u'^2\zeta}(\tau)$ are outside the confidence bounds. This combination strongly suggests that there is room for nonlinear terms to be added to the model.

4.3.2) Best nonlinear model

The ERLS algorithm with a step-wise regression procedure was used to estimate the coefficients in the nonlinear model. The initial specification was

The order of the lagged output,	$n_y = 3$
The order of the lagged input,	$n_u = 3$
The order of the noise,	$n_e = 3$
Time delay of the input,	TD = 5
Degree of nonlinearity of the model,	$l = 2$
Degree of nonlinearity of the noise,	$l_e = 1$

With this initial specification the total number of possible terms in the model was 55. Allowing the step-wise regression algorithm to sort through all possible terms using the Fisher-ratio test operating with 95% confidence bounds produced a MVT

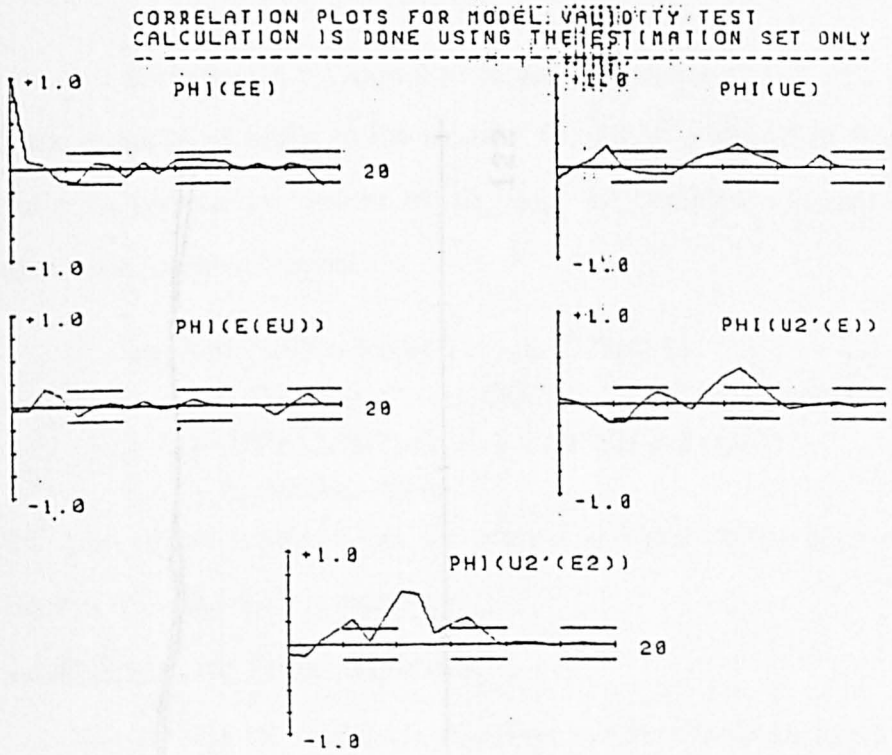


Figure 4.10: Model validity tests improved - best linear

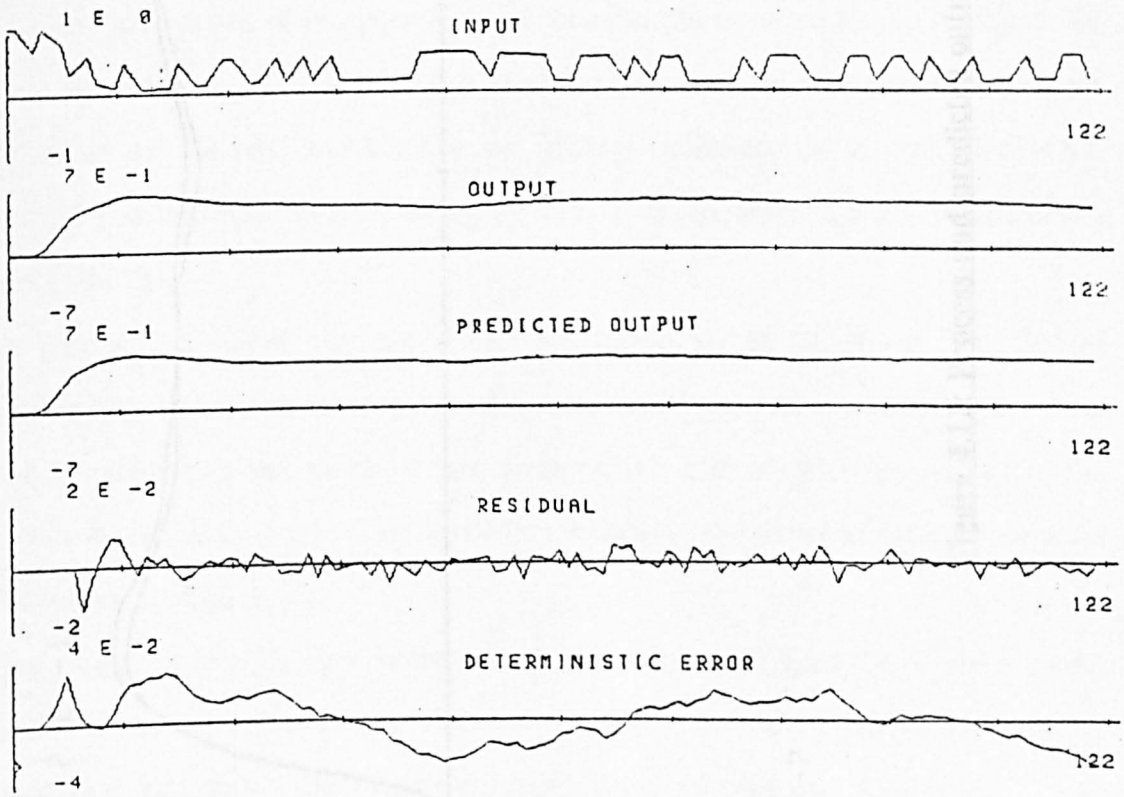


Figure 4.11: Plot of the residuals - best linear

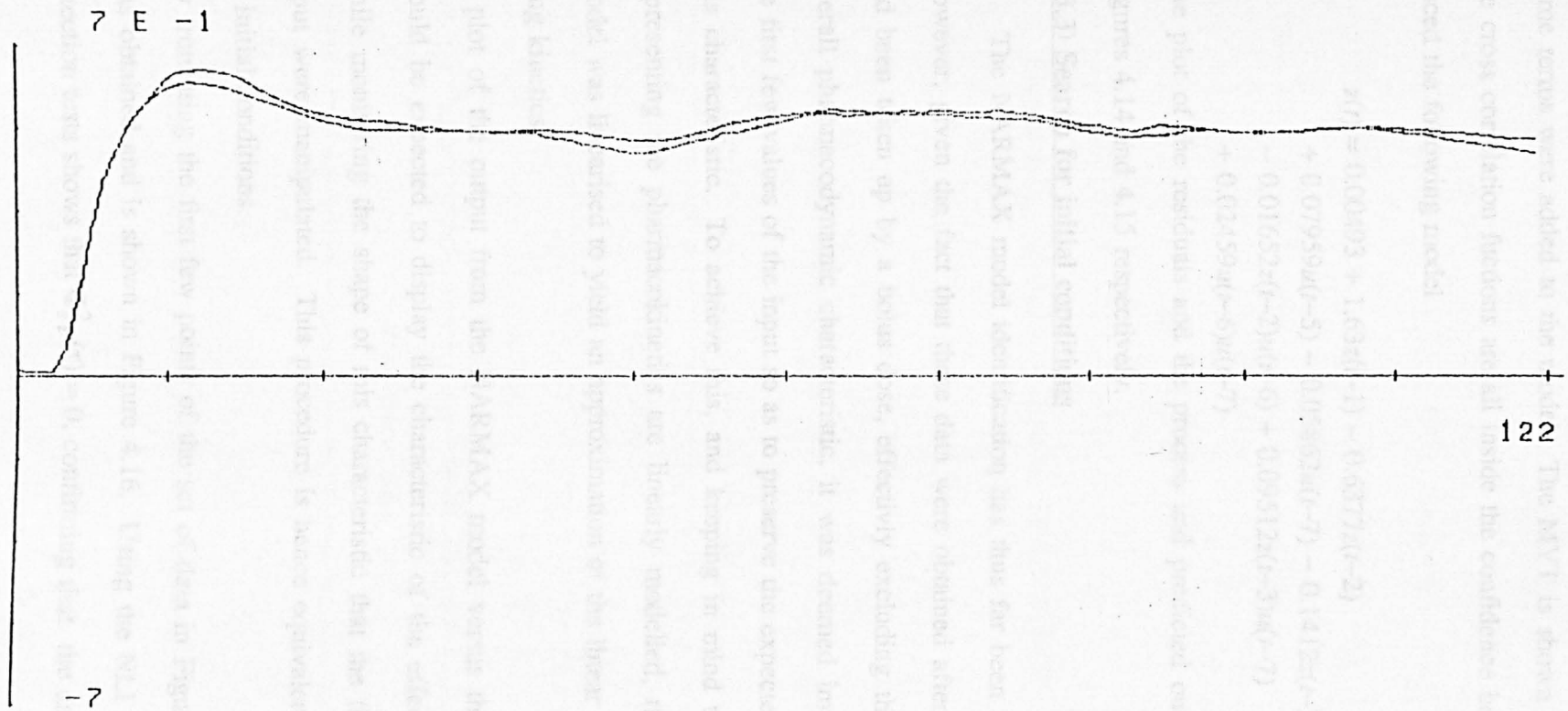


Figure 4.12: Process and predicted output superimposed - best linear

which at first showed the model to be deficient in some way. To improve on this, some terms were added to the model. The MVT is shown in Figure 4.13, where the cross correlation fuctions are all inside the confidence bounds, and which produced the following model

$$\begin{aligned} z(t) = {} & 0.00493 + 1.63z(t-1) - 0.6377z(t-2) \\ & + 0.07959u(t-5) - 0.05862u(t-7) - 0.1412z(t-3)u(t-5) \\ & - 0.01652z(t-2)u(t-6) + 0.09512z(t-3)u(t-7) \\ & + 0.02459u(t-6)u(t-7) \end{aligned} \tag{4.3}$$

The plot of the residuals and the process and predicted outputs are displayed in Figures 4.14 and 4.15 respectively.

4.3.3) Search for initial conditions

The NARMAX model identification has thus far been fitted to the raw data. However, given the fact that these data were obtained after the margin of safety had been taken up by a bolus dose, effectivly excluding the dead-zone from the overall pharmacodynamic characteristic, it was deemed important to manipulate the first few values of the input so as to preserve the expected known ' shape ' of this characteristic. To achieve this, and keeping in mind that the compartments representing the pharmacokinetics are linearly modelled, the overall NARMAX model was linearised to yield an approximation of the linear dynamic model of the drug kinetics.

A plot of the output from the NARMAX model versus the above linear model would be expected to display the characteristic of the effect compartment. It is while monitoring the shape of this characteristic that the first few values of the input were manipulated. This procedure is hence equivalent to introducing a set of initial conditions.

By truncating the first few points of the set of data in Figure 4.4, a linear model was obtained and is shown in Figure 4.16. Using the NLI package, the structure detection tests shows that $\Phi^2_{z'z'}(\tau) = 0$, confirming that the data were linear

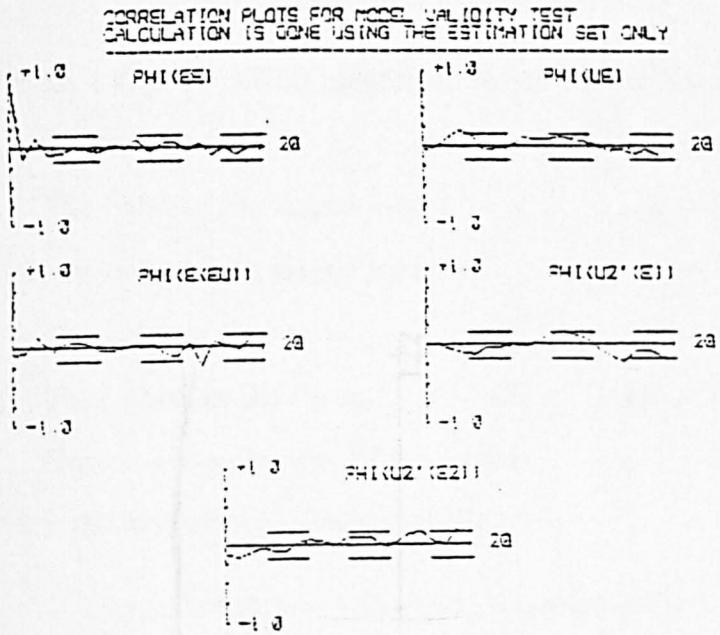


Figure 4.13: Model validity tests - best nonlinear

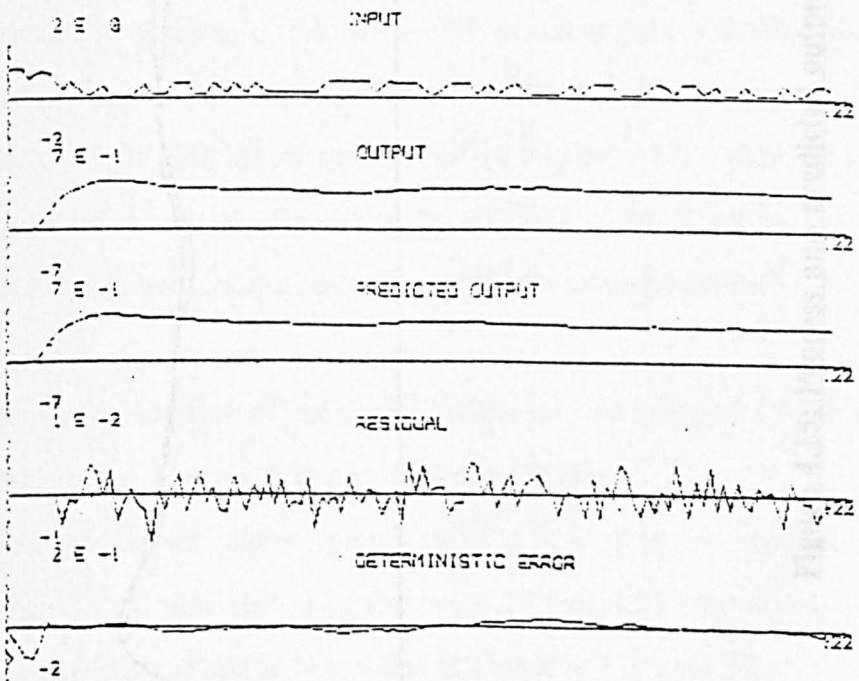


Figure 4.14: Plot of the residuals - best nonlinear

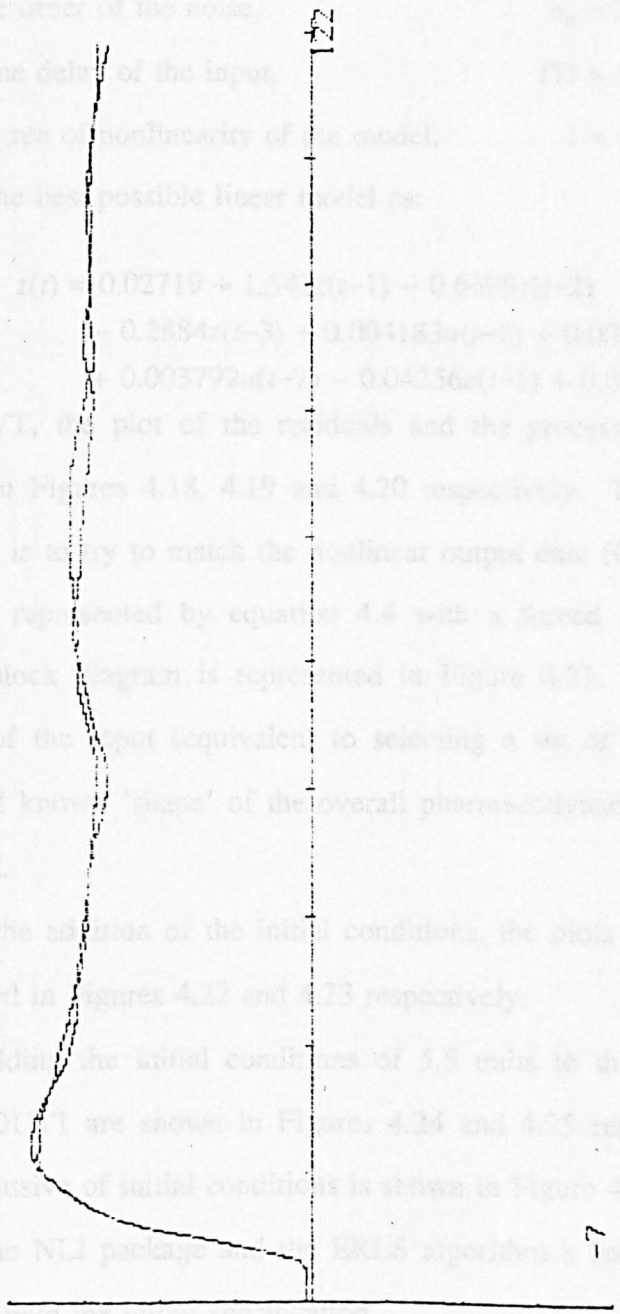


Figure 4.15: Process and predicted output superimposed - best nonlinear

(Figure 4.17). An ERLS algorithm was used to estimate the coefficients. The initial specification is

The order of the lagged output, $n_y = 3$

The order of the lagged input, $n_u = 3$

The order of the noise, $n_e = 2$

Time delay of the input, TD = 5

Degree of nonlinearity of the model, 1 = 1

giving the best possible linear model as:

$$z(t) = 0.02719 + 1.541z(t-1) - 0.6699z(t-2) - 0.2884z(t-3) + 0.004183u(t-5) + 0.001354u(t-6) + 0.003792u(t-7) - 0.04256e(t-1) + 0.02517e(t-2) \quad (4.4)$$

The MVT, the plot of the residuals and the process and predicted outputs are shown in Figures 4.18, 4.19 and 4.20 respectively. The aim of the 'phase plot' method is to try to match the nonlinear output data (OUT2) and the linear model (OUT1) represented by equation 4.4 with a forced nonlinear input data. The whole block diagram is represented in Figure 4.21. Manipulating the first few values of the input (equivalent to selecting a set of initial conditions) until the expected known 'shape' of the overall pharmacodynamics (Hill equation type) is obtained.

Before the addition of the initial conditions, the plots of OUT2 versus OUT1 are illustrated in Figures 4.22 and 4.23 respectively.

After adding the initial conditions of 5.5 units to the input the plots of OUT2 versus OUT1 are shown in Figures 4.24 and 4.25 respectively. The new set of data inclusive of initial conditions is shown in Figure 4.26.

Using the NLI package and the ERLS algorithm a best nonlinear fit is found for the data with the initial specification

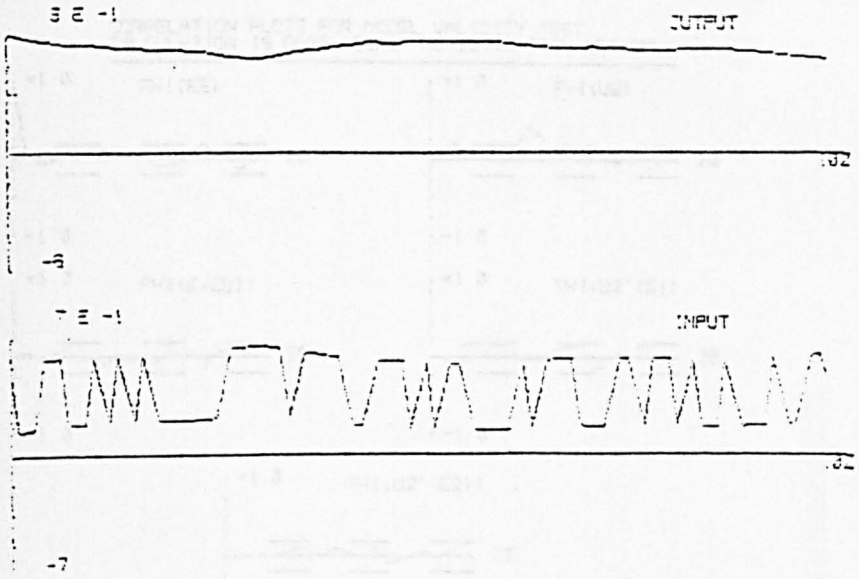


Figure 4.16: Representation of the truncated data

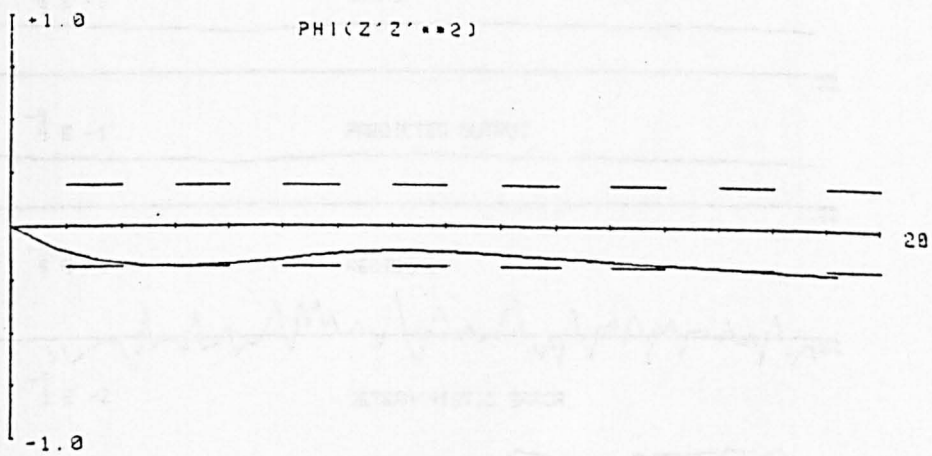


Figure 4.17: Structure detection test of the truncated data

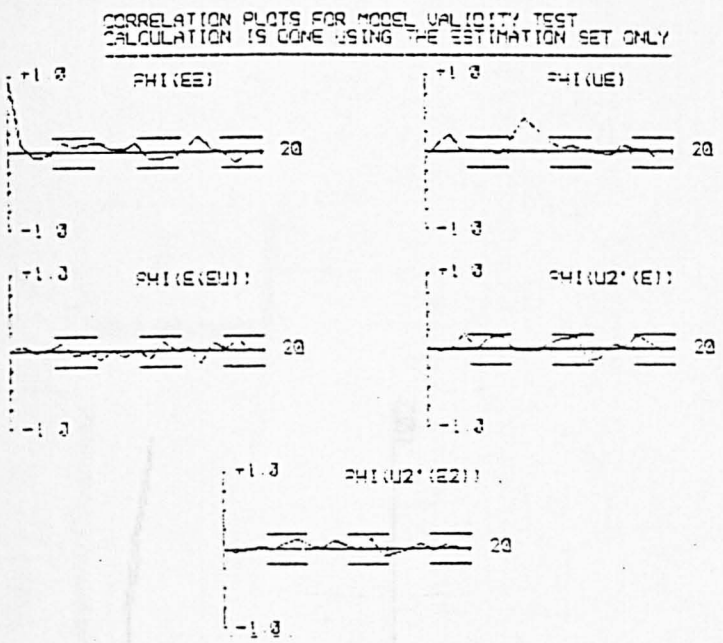


Figure 4.18: Model validity tests of the truncated data

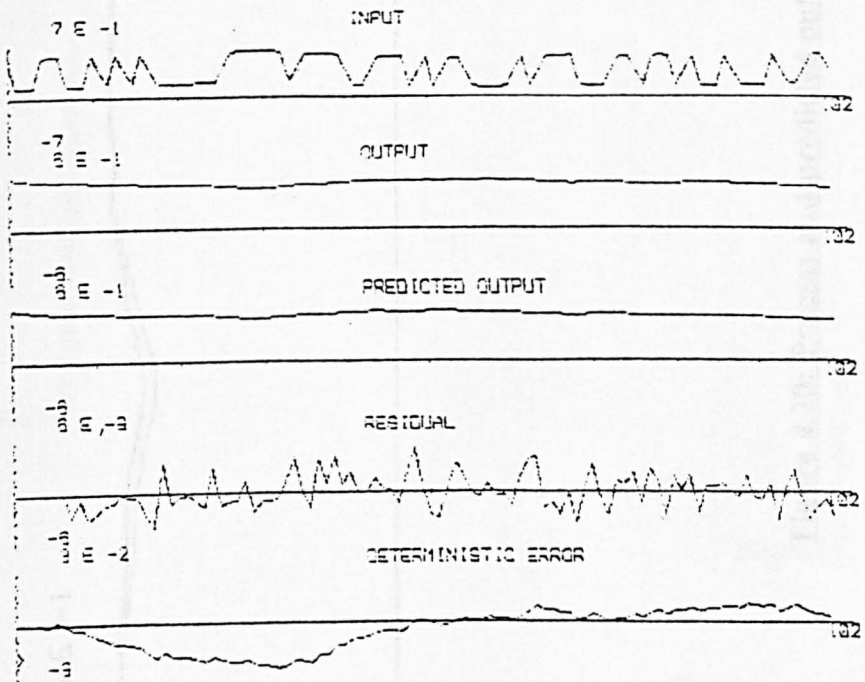


Figure 4.19: Plot of the residuals of the truncated data

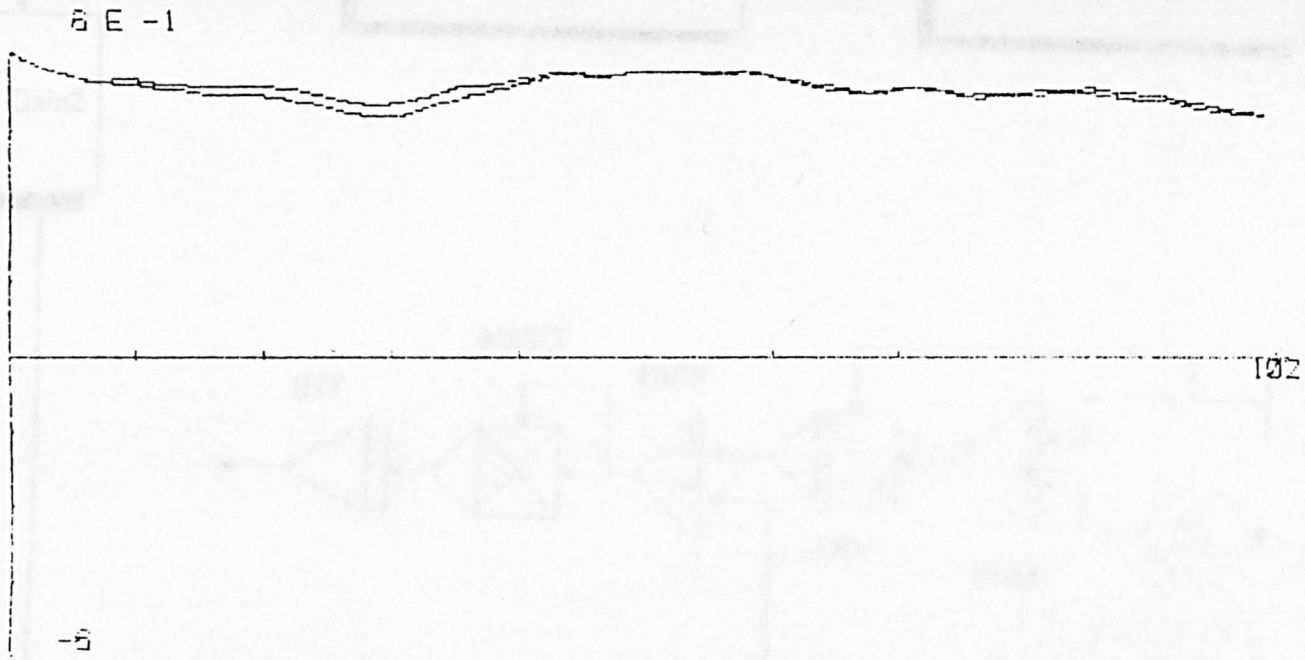


Figure 4.20: Process and predicted output superimposed of the truncated data

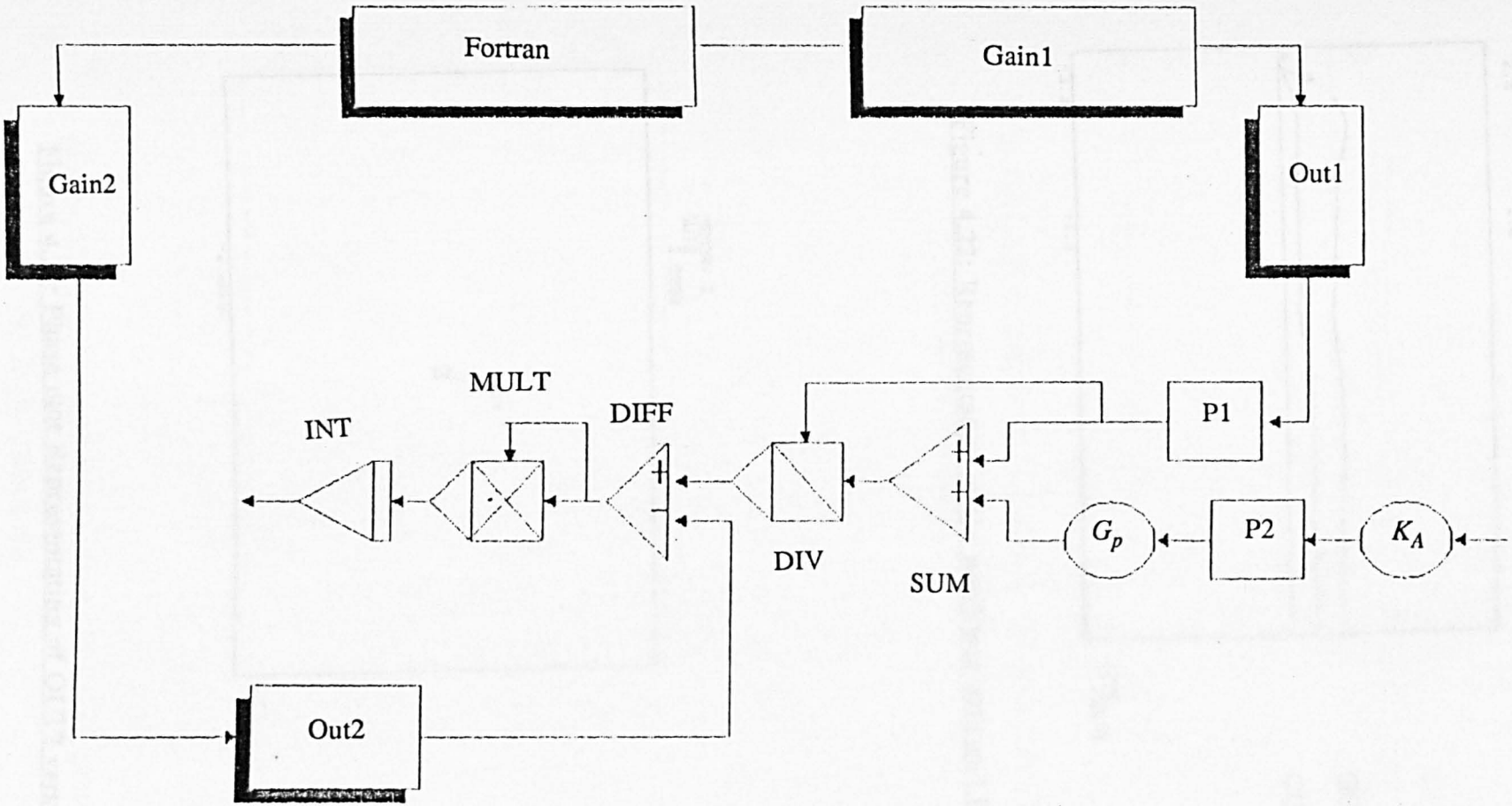
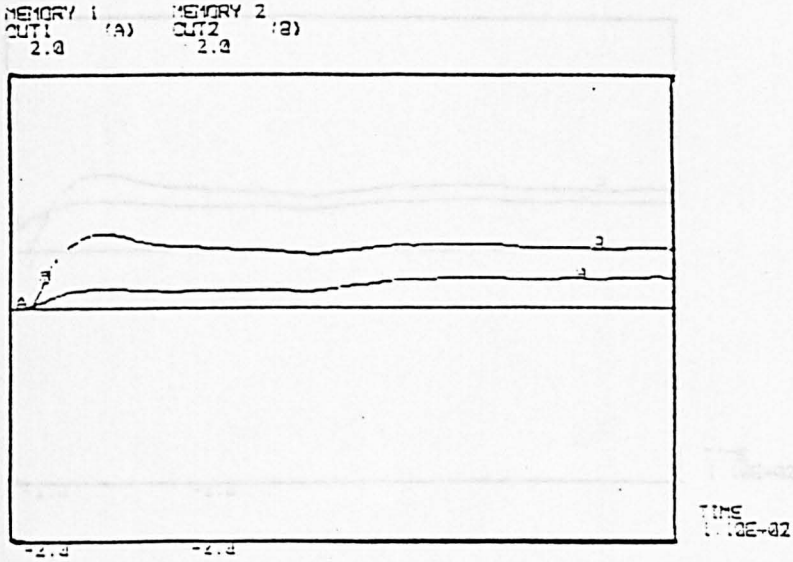


Figure 4.21: Block diagram of the phase plot.



OUT2 - nonlinear data

OUT1 - linear model

Figure 4.22: Representation of the nonlinear data and linear model

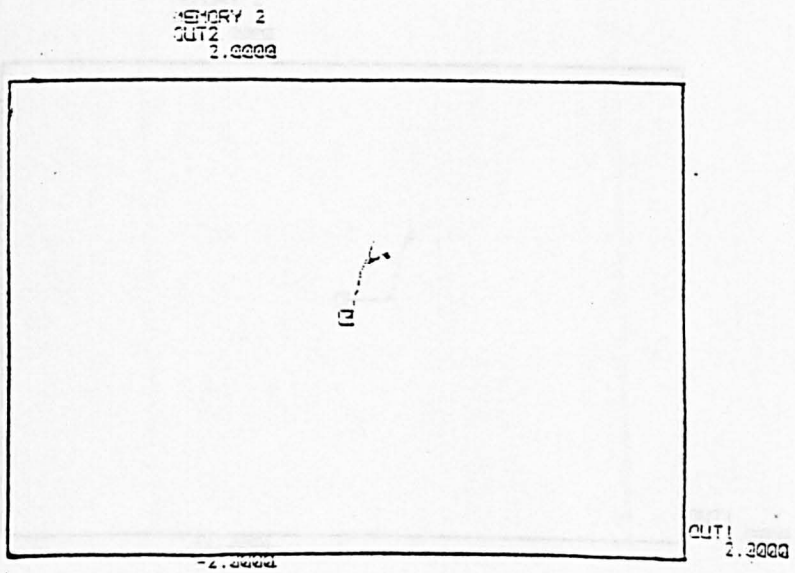
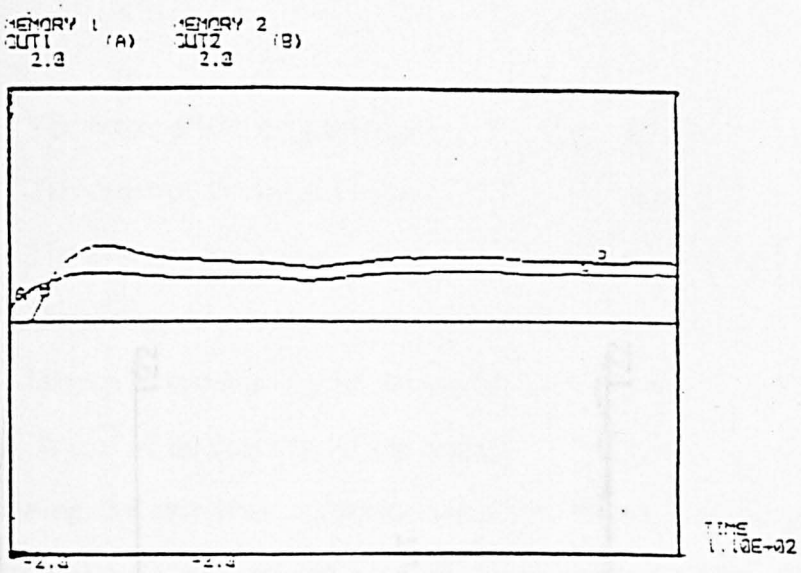


Figure 4.23: Phase plot Representating of OUT2 versus OUT1



OUT2 - nonlinear data

OUT1 - linear model

Figure 4.24: Representation of the nonlinear data and linear model with initial conditions

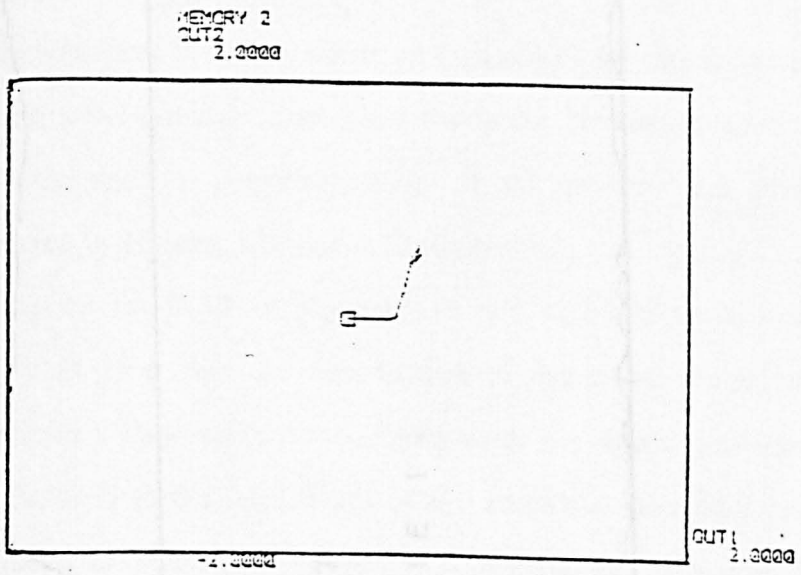


Figure 4.25: Phase plot Representating of OUT2 versus OUT1 with initial conditions

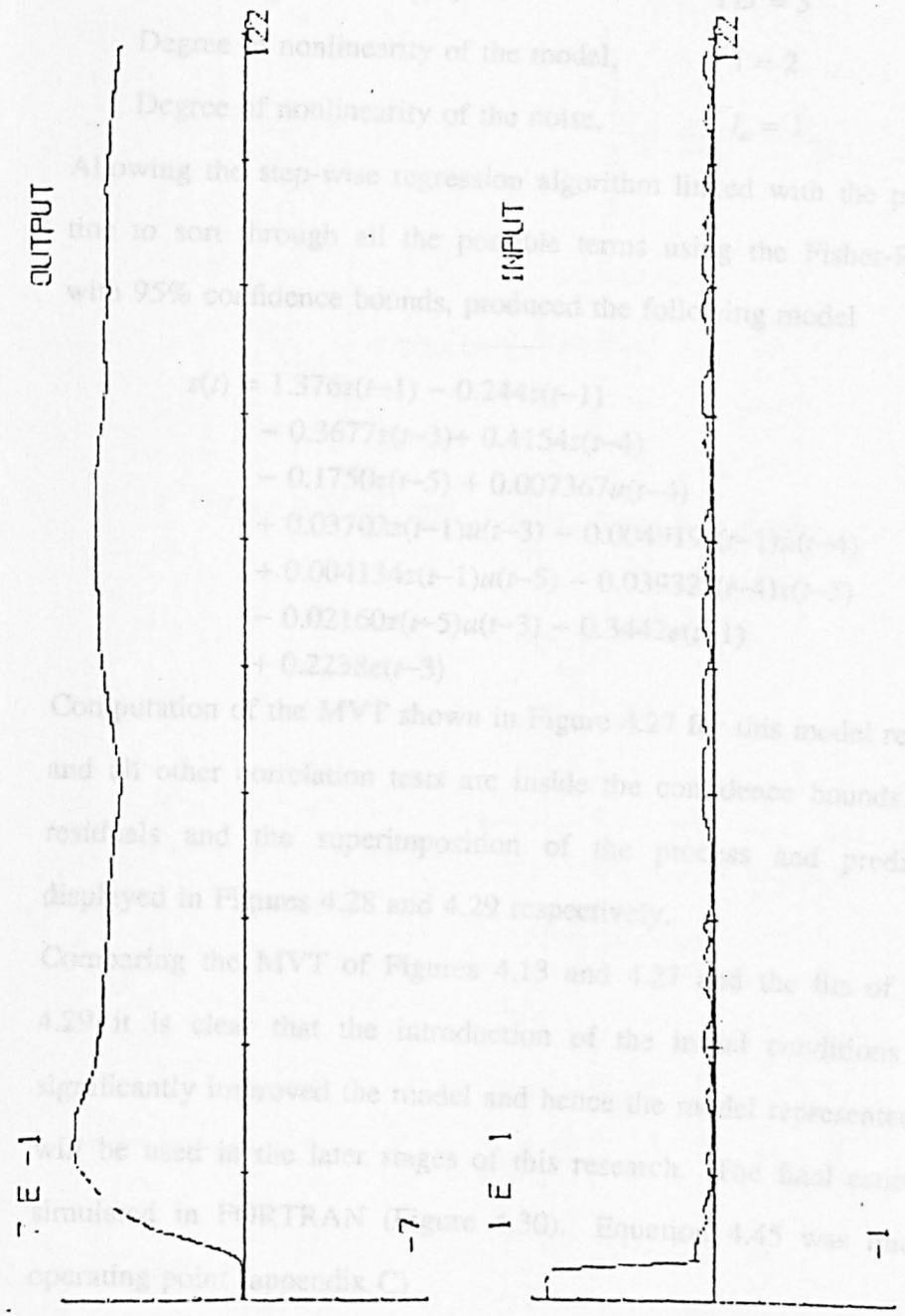


Figure 4.26: Representation of the data with initial conditions

The order of the lagged output, $n_y = 5$

The order of the lagged input, $n_u = 3$

The order of the noise, $n_e = 5$

Time delay of the input, TD = 3

Degree of nonlinearity of the model, l = 2

Degree of nonlinearity of the noise, $l_e = 1$

Allowing the step-wise regression algorithm linked with the prediction error routine to sort through all the possible terms using the Fisher-Ratio test operating with 95% confidence bounds, produced the following model

$$\begin{aligned} z(t) = {} & 1.376z(t-1) - 0.244z(t-1) \\ & - 0.3677z(t-3) + 0.4154z(t-4) \\ & - 0.1750z(t-5) + 0.007367u(t-4) \\ & + 0.03702z(t-1)u(t-3) - 0.004919z(t-1)u(t-4) \\ & + 0.004134z(t-1)u(t-5) - 0.03932z(t-4)z(t-5) \\ & - 0.02160z(t-5)u(t-3) - 0.3442e(t-1) \\ & + 0.2238e(t-3) \end{aligned} \tag{4.5}$$

Computation of the MVT shown in Figure 4.27 for this model revealed that $\Phi_{\zeta\zeta}(\tau)$ and all other correlation tests are inside the confidence bounds. The plot of the residuals and the superimposition of the process and predicted outputs are displayed in Figures 4.28 and 4.29 respectively.

Comparing the MVT of Figures 4.13 and 4.27 and the fits of Figures 4.15 and 4.29 it is clear that the introduction of the initial conditions to the data has significantly improved the model and hence the model represented by equation 4.5 will be used in the later stages of this research. The final estimated model was simulated in FORTRAN (Figure 4.30). Equation 4.45 was linearized about an operating point (appendix C)

$$\begin{aligned} z_L(t) = {} & 8.0288210^{-3} + 1.389z(t-1) - 0.244z(t-2) \\ & - 0.3627z(t-3) + 0.394z(t-4) - 0.2037z(t-5) \\ & + 8.11310^{-}3u(t-3) + 4.778810^{-3}u(t-4) + 2.17510^{-3}u(t-5) \end{aligned}$$

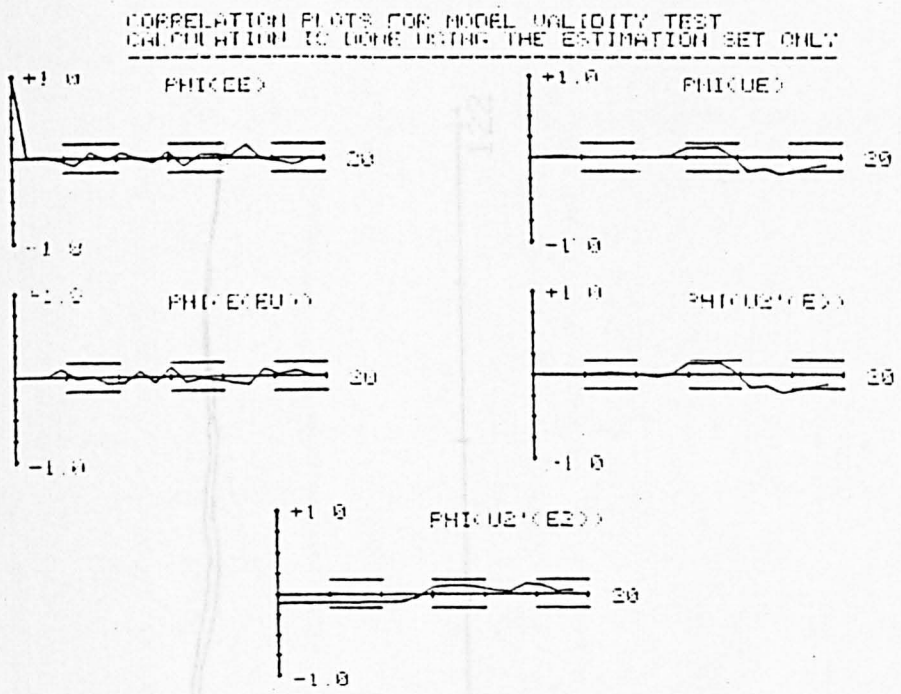


Figure 4.27: Model validity tests - best nonlinear

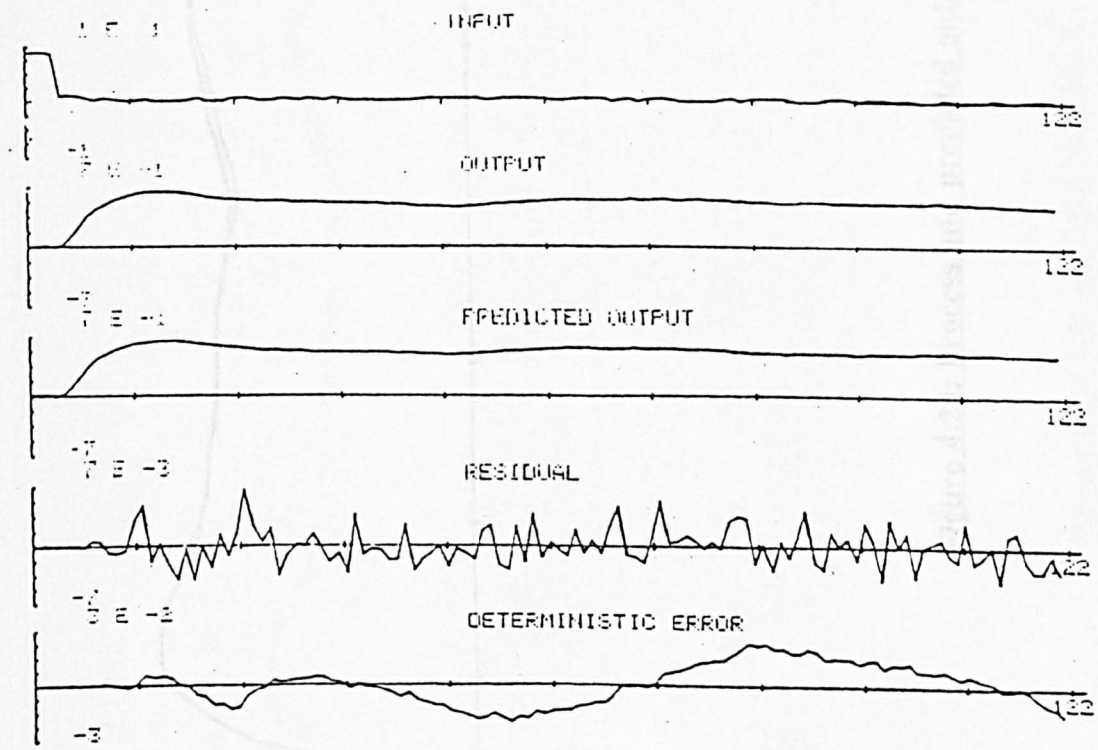


Figure 4.28: Plot of the residuals - best nonlinear

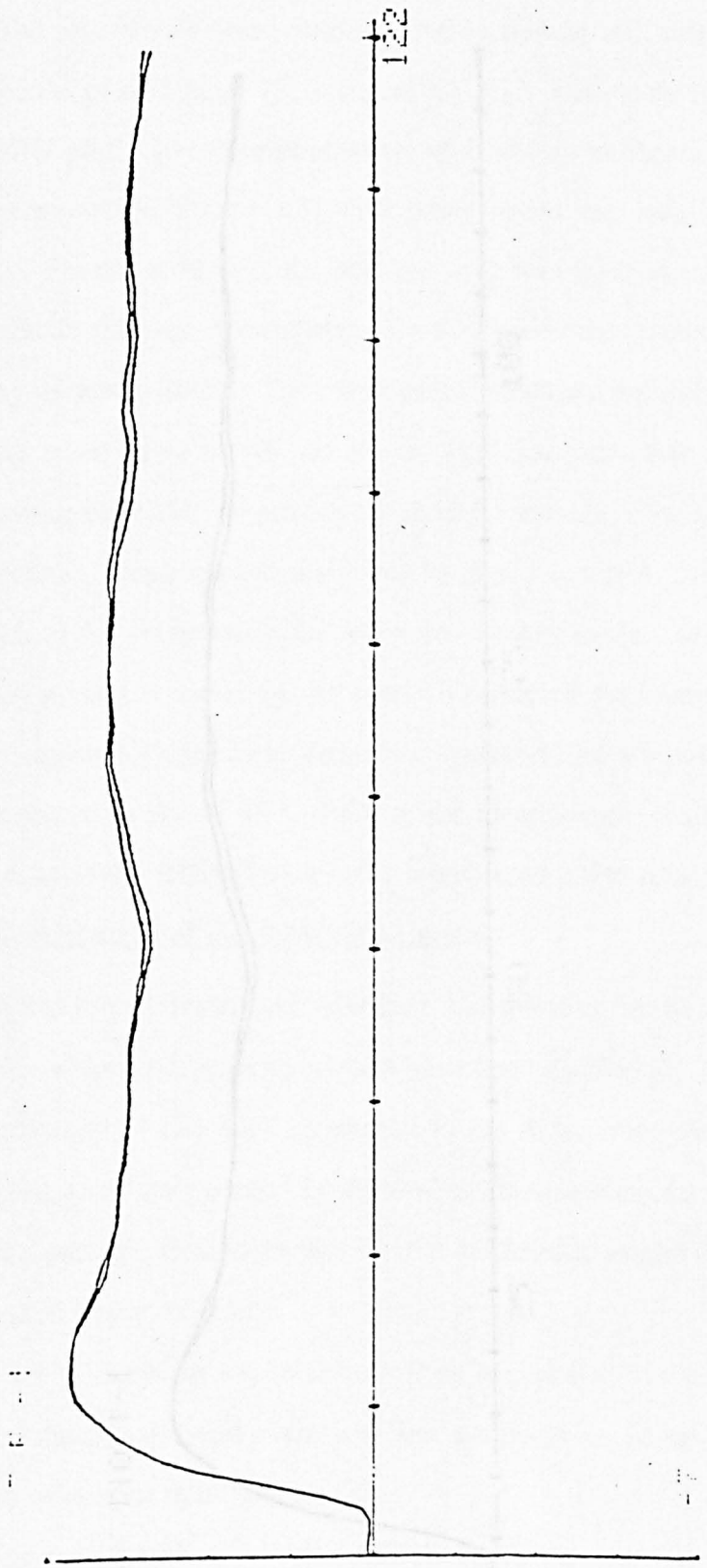


Figure 4.22: Process and predicted output superimposed - best nonlinear

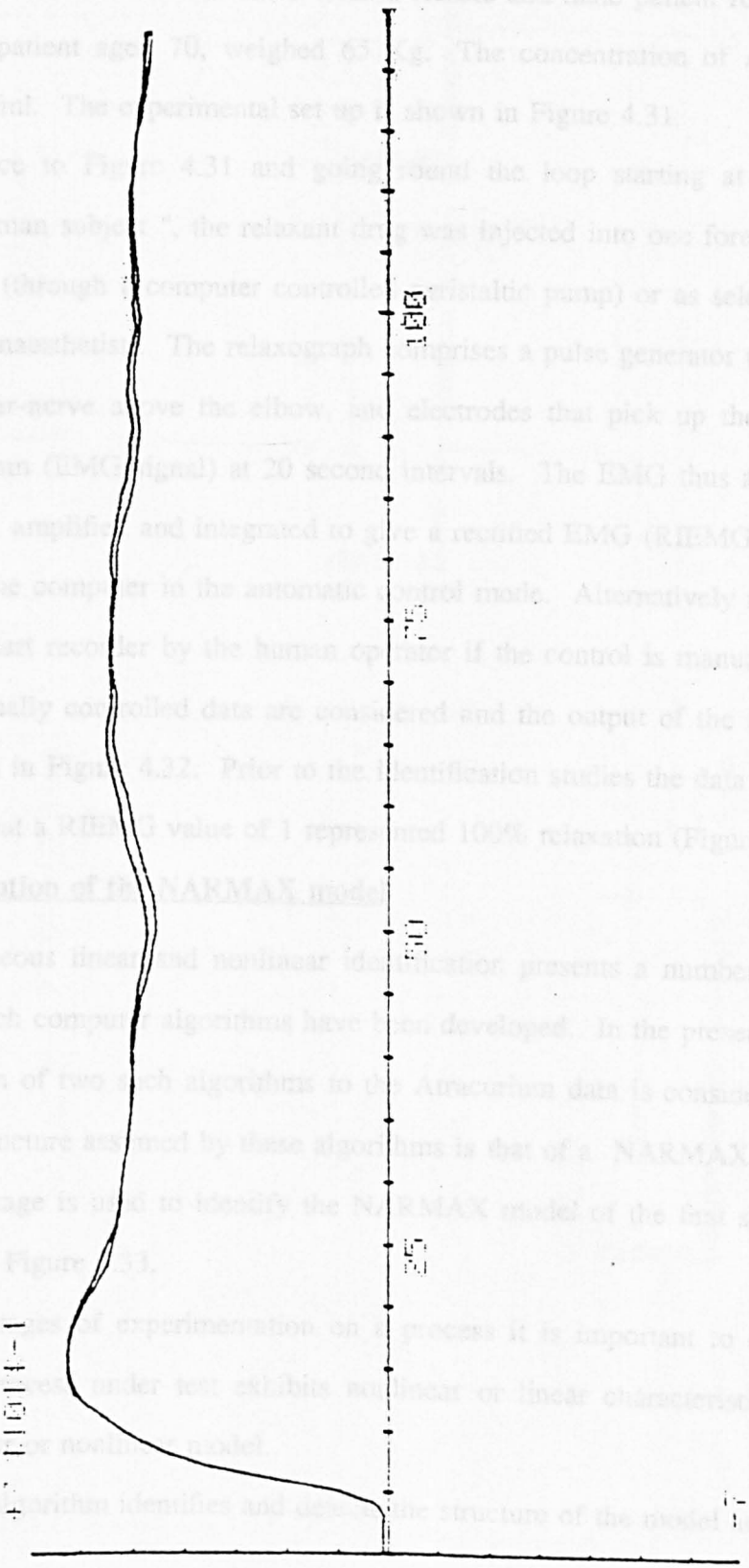


Figure 4.30: Check of the NARMAX model

4.4) Data acquisition and conditioning for Atracurium

Two sets of data were obtained from a female and male patient respectively. The female patient aged 70, weighed 65 Kg. The concentration of Atracurium was 500 μ g/ml. The experimental set up is shown in Figure 4.31.

With reference to Figure 4.31 and going round the loop starting at the block named, " Human subject ", the relaxant drug was injected into one forearm either continuously (through a computer controlled peristaltic pump) or as selected infusion (by an anaesthetist). The relaxograph comprises a pulse generator that stimulates the ulnar-nerve above the elbow, and electrodes that pick up the resulting electromyogram (EMG signal) at 20 second intervals. The EMG thus acquired is then rectified, amplified and integrated to give a rectified EMG (RIEMG) which is fed back to the computer in the automatic control mode. Alternatively it is monitored on a chart recorder by the human operator if the control is manual. In this chapter, manually controlled data are considered and the output of the first set of data is shown in Figure 4.32. Prior to the identification studies the data were normalised, so that a RIEMG value of 1 represented 100% relaxation (Figure 4.33).

4.5) Identification of the NARMAX model

Simultaneous linear and nonlinear identification presents a number of problems, for which computer algorithms have been developed. In the present section, the application of two such algorithms to the Atracurium data is considered. The underlying structure assumed by these algorithms is that of a NARMAX model.

The NLI package is used to identify the NARMAX model of the first set of data represented in Figure 4.33.

In the early stages of experimentation on a process it is important to determine whether the process under test exhibits nonlinear or linear characteristics which warrant a linear or nonlinear model.

Although the algorithm identifies and detects the structure of the model and then

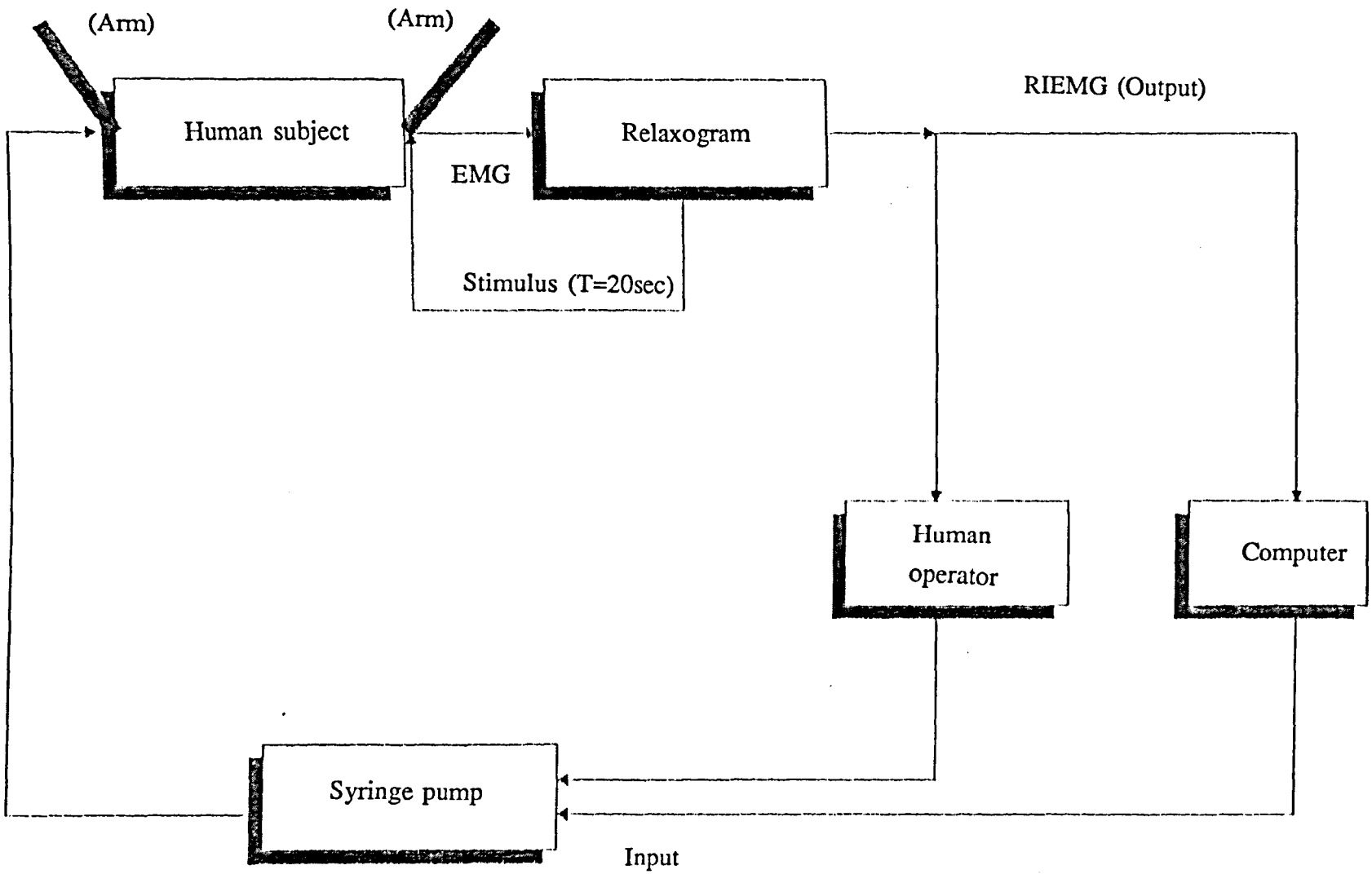


Figure 4.31: Schematic of experimental procedure for muscle relaxant administration of Atracurium

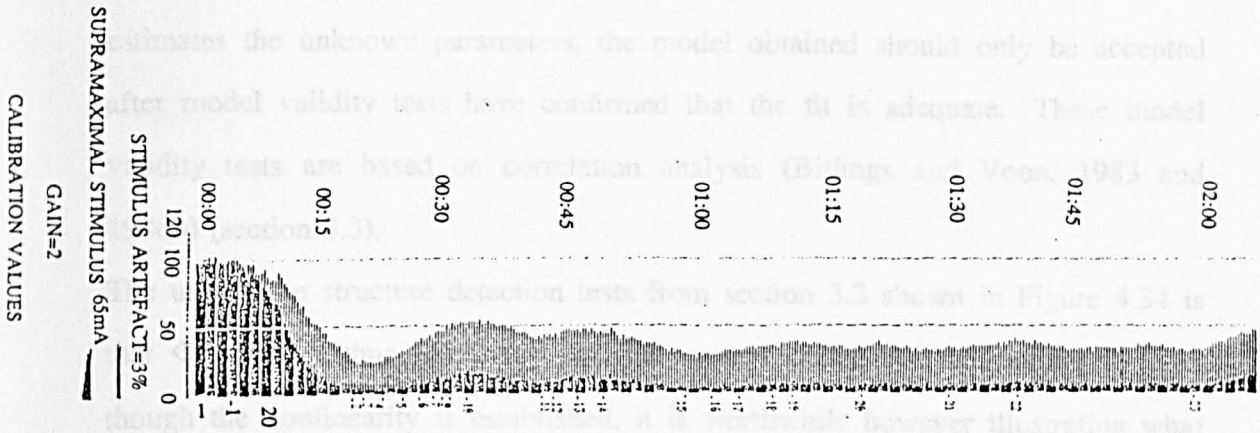


Figure 4.32: Recorded E.M.G obtained during an operation performed by an anaesthetist with a patient aged 70 years. Desired level: 20%; drug: Atracurium.

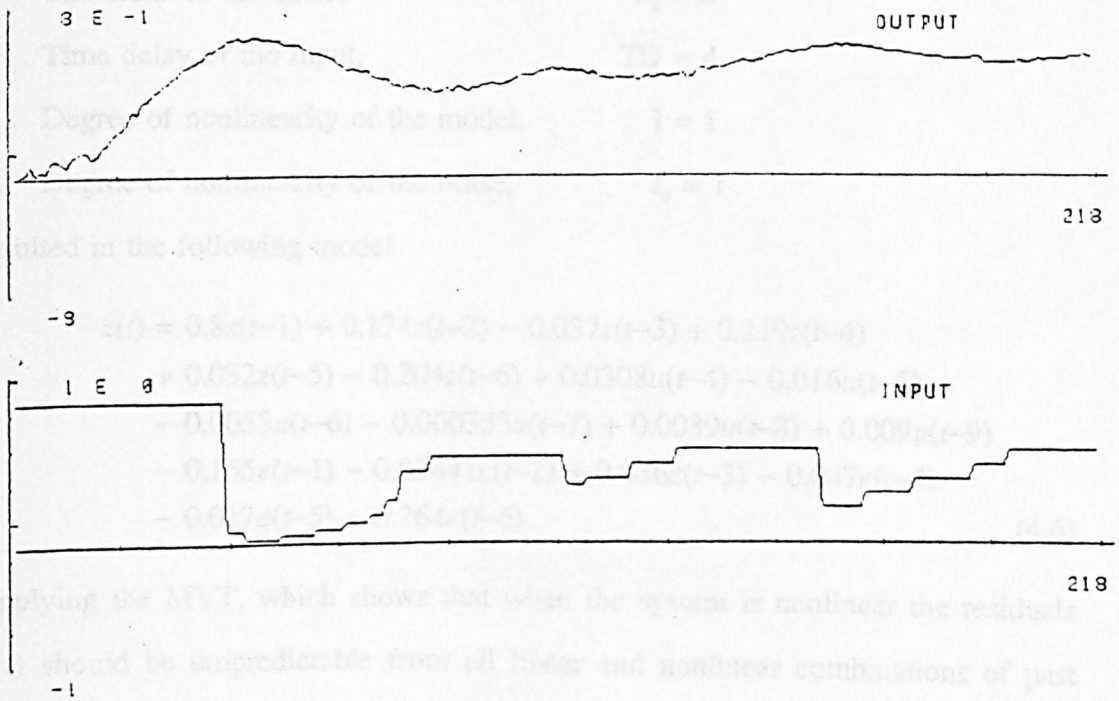


Figure 4.33: RIEMG and infusion data of Atracurium

estimates the unknown parameters, the model obtained should only be accepted after model validity tests have confirmed that the fit is adequate. These model validity tests are based on correlation analysis (Billings and Voon, 1983 and 1986b) (section 3.3).

The use of the structure detection tests from section 3.3 shown in Figure 4.34 is that $\Phi^2_{z'z'}(\tau)\neq 0$, thus indicating that the system under test is nonlinear. Even though the nonlinearity is established, it is worthwhile however illustrating what happens if a linear model is fitted to the data.

4.5.1) Best linear model fit

Extensive preliminary analysis of the data involved estimating the coefficients in linear models of varying process and noise model orders and time delays. Using a recursive extended least squares with the following specification

The order of the lagged output, $n_y = 6$

The order of the lagged input, $n_u = 6$

The order of the noise, $n_e = 6$

Time delay of the input, TD = 4

Degree of nonlinearity of the model, $l = 1$

Degree of nonlinearity of the noise, $l_e = 1$

resulted in the following model

$$\begin{aligned} z(t) = {} & 0.8z(t-1) + 0.174z(t-2) - 0.037z(t-3) + 0.219z(t-4) \\ & + 0.032z(t-5) - 0.204z(t-6) + 0.0308u(t-4) - 0.016u(t-5) \\ & - 0.0058u(t-6) - 0.000353u(t-7) + 0.0039u(t-8) + 0.009u(t-9) \\ & - 0.155e(t-1) - 0.03441e(t-2) + 0.036e(t-3) - 0.047e(t-4) \\ & - 0.017e(t-5) + 0.264e(t-6) \end{aligned} \tag{4.6}$$

Applying the MVT, which shows that when the system is nonlinear the residuals $\zeta(\tau)$ should be umpredictable from all linear and nonlinear combinations of past inputs and outputs (section 3.3.2).

The MVT is shown in Figure 4.35. From Figure 4.35 the correlation tests $\Phi_{\zeta\zeta}(\tau)$

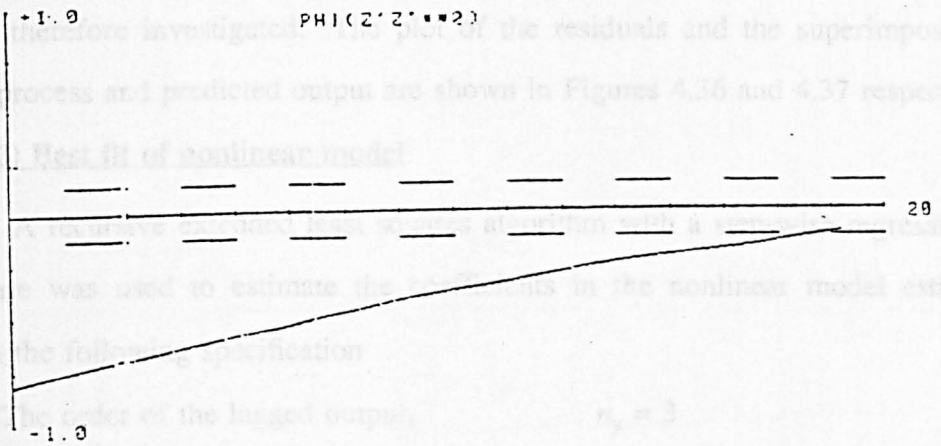


Figure 4.34: Structure detection test of the actual data

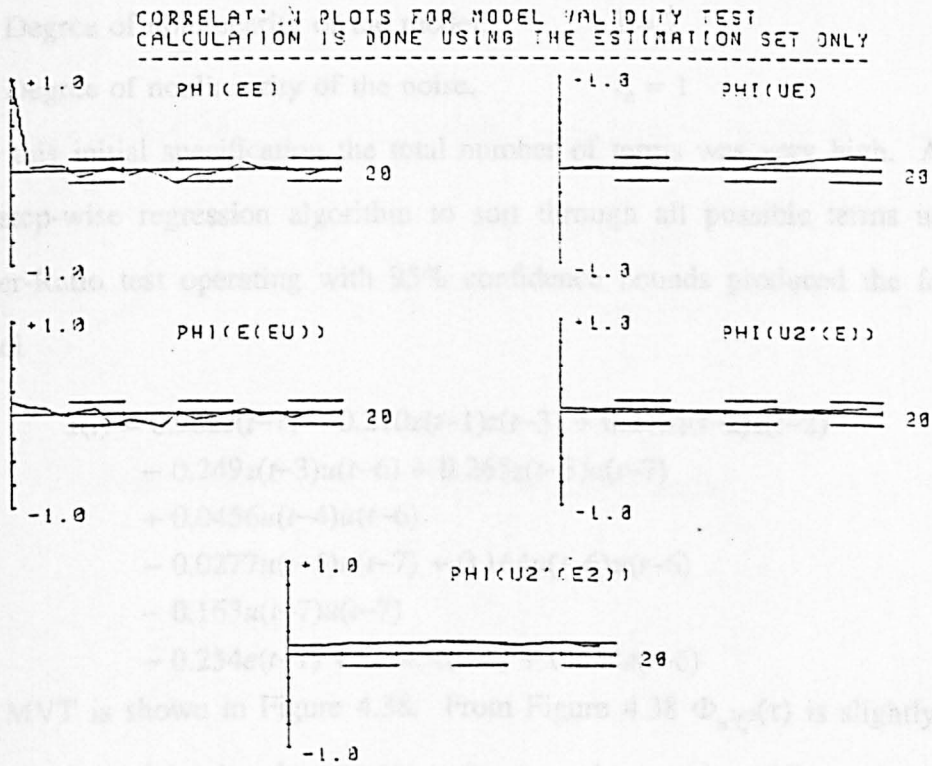


Figure 4.35: Model validity tests - best linear

and $\Phi_{u\zeta}(\tau)$ are well inside the 95% confidence bounds, but $\Phi_{\zeta(\zeta u)}(\tau)$ and $\Phi_{u^2\zeta^2}(\tau)$ are slightly outside the confidence bounds. These combinations suggest that non-linear terms are missing. The introduction of these nonlinear terms into the model was therefore investigated. The plot of the residuals and the superimposition of the process and predicted output are shown in Figures 4.36 and 4.37 respectively.

4.5.2) Best fit of nonlinear model

A recursive extended least squares algorithm with a step-wise regression procedure was used to estimate the coefficients in the nonlinear model estimation, with the following specification

The order of the lagged output,	$n_y = 3$
the order of the lagged input,	$n_u = 4$
The order of the noise,	$n_e = 6$
Time delay of the input,	TD = 4
Degree of nonlinearity of the model,	l = 2
Degree of nonlinearity of the noise,	$l_e = 1$

with this initial specification the total number of terms was very high. Allowing the step-wise regression algorithm to sort through all possible terms using the Fisher-Ratio test operating with 95% confidence bounds produced the following model

$$\begin{aligned} z(t) = {} & 0.982z(t-1) - 0.210z(t-1)z(t-3) + 0.213z(t-2)z(t-2) \\ & - 0.249z(t-3)u(t-6) + 0.265z(t-3)u(t-7) \\ & + 0.0456u(t-4)u(t-6) \\ & - 0.0277u(t-5)u(t-7) + 0.164u(t-6)u(t-6) \\ & - 0.163u(t-7)u(t-7) \\ & - 0.254e(t-1) + 0.145e(t-4) + 0.224e(t-6) \end{aligned} \tag{4.7}$$

The MVT is shown in Figure 4.38. From Figure 4.38 $\Phi_{u^2\zeta^2}(\tau)$ is slightly outside the confidence bounds. An attempt to improve the tests by adding a few terms to the model did not alter the fit, and hence the best nonlinear model obtained is

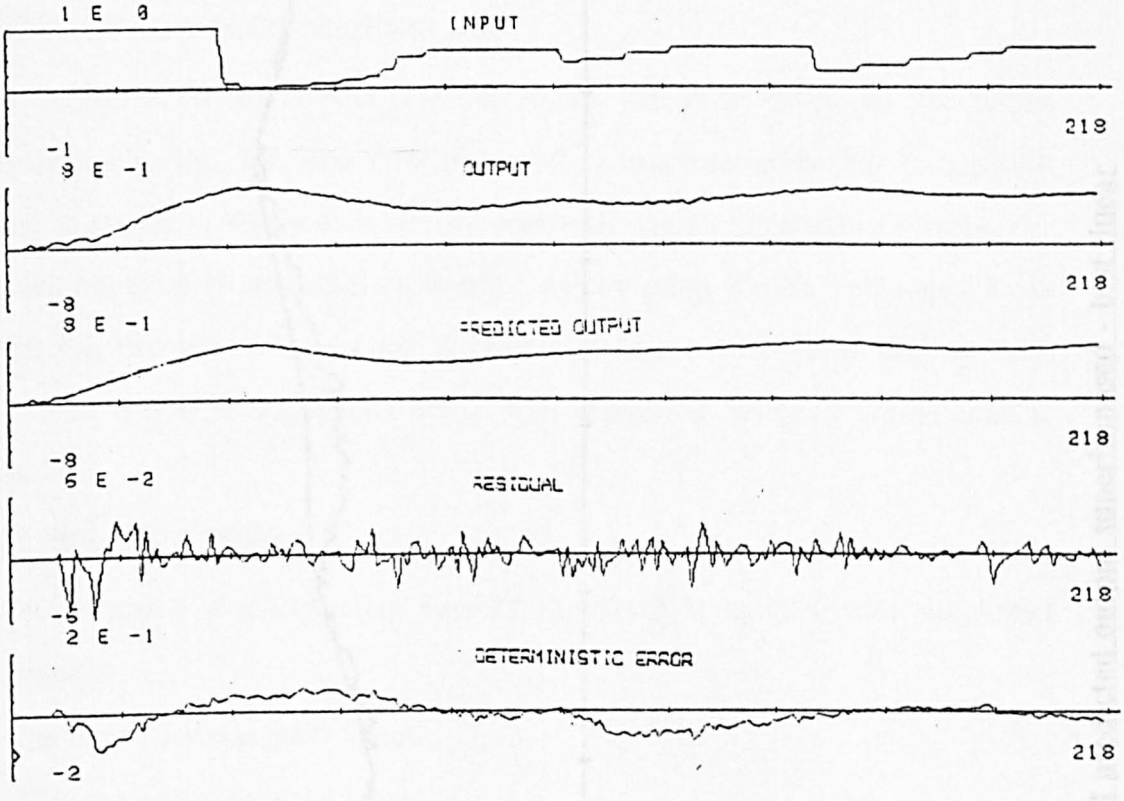


Figure 4.36: Plot of the residuals - best linear

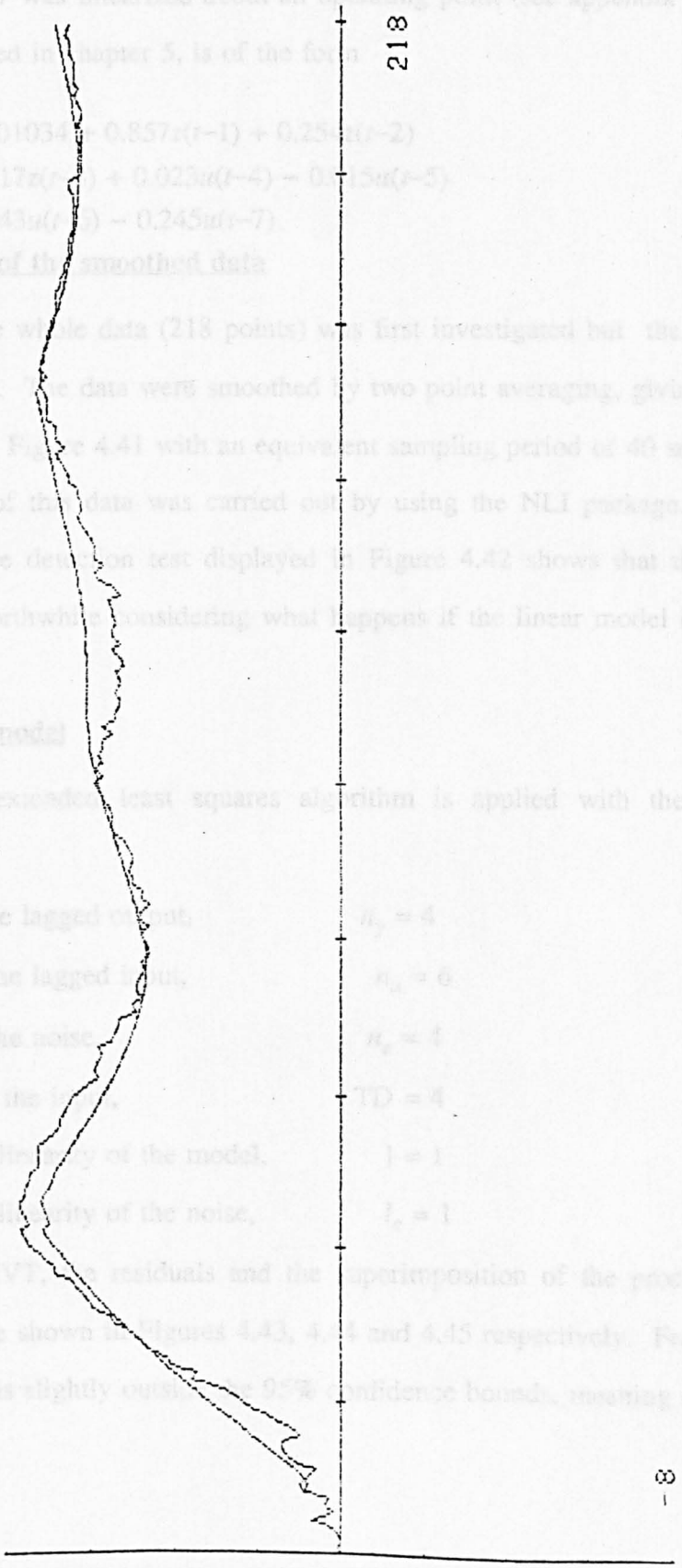


Figure 4.37: Process and predicted output superimposed - best linear

represented by equation 4.7. The plot of the residuals and the superimposition of the process and predicted outputs are displayed in Figures 4.39 and 4.40 respectively. Equation 4.7 was linearized about an operating point (see appendix C), the resulting model used in chapter 5, is of the form

$$\begin{aligned} z_l(t) = &-0.01034 + 0.857z(t-1) + 0.254z(t-2) \\ &- 0.117z(t-3) + 0.023u(t-4) - 0.015u(t-5) \\ &+ 0.043u(t-6) - 0.245u(t-7) \end{aligned} \tag{4.8}$$

4.6) Identification of the smoothed data

The use of the whole data (218 points) was first investigated but the results were unsatisfactory. The data were smoothed by two point averaging, giving (108 points) as shown in Figure 4.41 with an equivalent sampling period of 40 seconds. The identification of this data was carried out by using the NLI package. Even though the structure detection test displayed in Figure 4.42 shows that the data was linear, it is worthwhile considering what happens if the linear model is fitted to the data.

4.6.1) Best linear model

A recursive extended least squares algorithm is applied with the initial specification

The order of the lagged output,	$n_y = 4$
The order of the lagged input,	$n_u = 6$
The order of the noise,	$n_e = 4$
Time delay of the input,	TD = 4
Degree of nonlinearity of the model,	l = 1
Degree of nonlinearity of the noise,	$l_e = 1$

The plot of the MVT, the residuals and the superimposition of the process and predicted output are shown in Figures 4.43, 4.44 and 4.45 respectively. From Figure 4.43, $\Phi_{\zeta\zeta(u)}(\tau)$ is slightly outside the 95% confidence bounds, meaning that

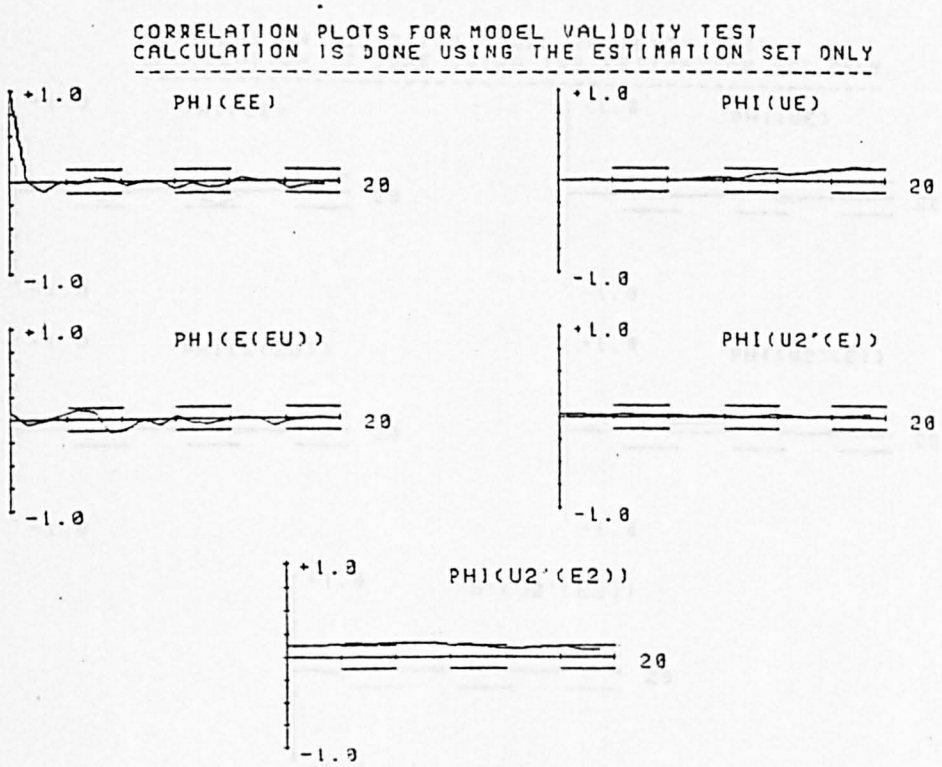


Figure 4.38: Model validity tests - best nonlinear

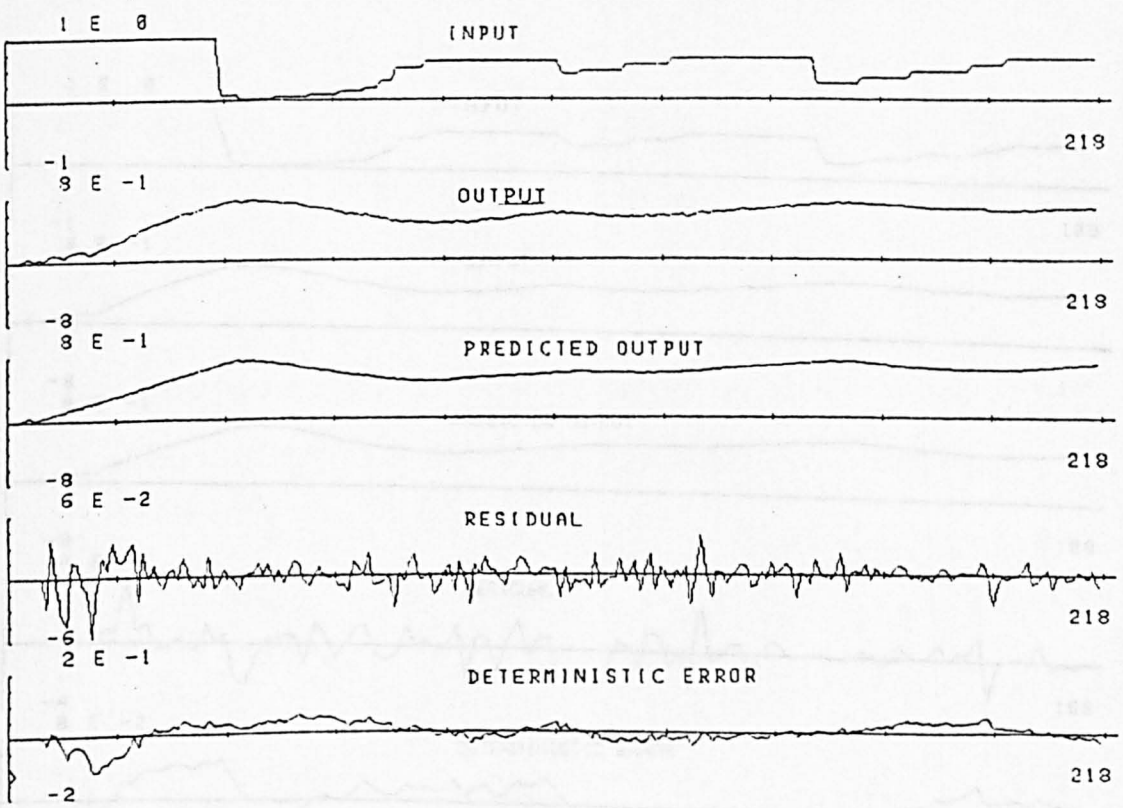


Figure 4.39: Plot of the residuals - best nonlinear

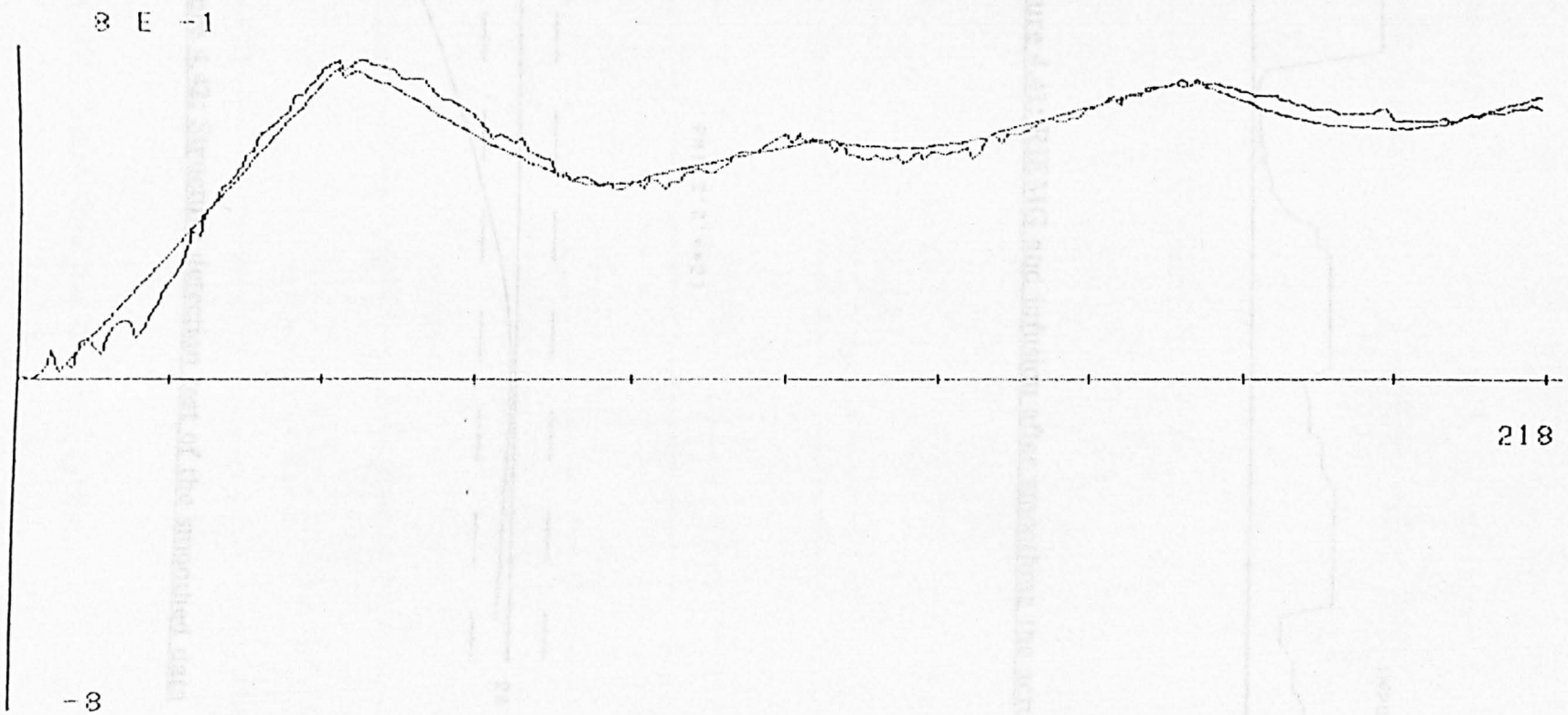


Figure 4.40: Process and predicted output superimposed - best nonlinear

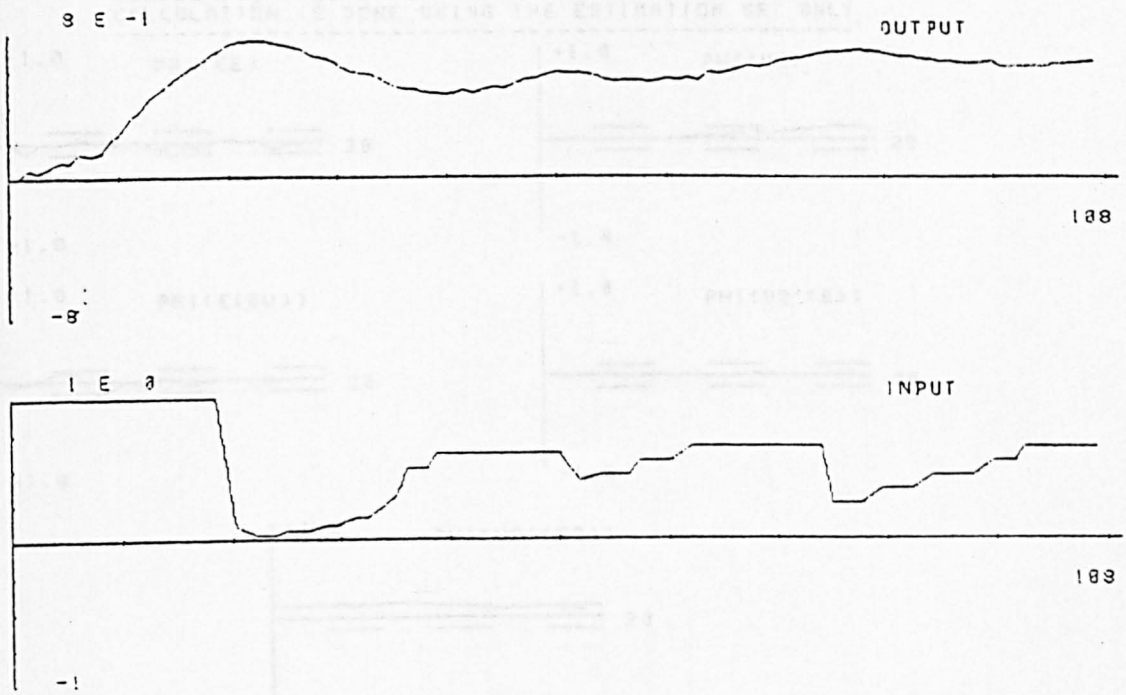


Figure 4.41: RIEMG and infusion after smoothing the actual data

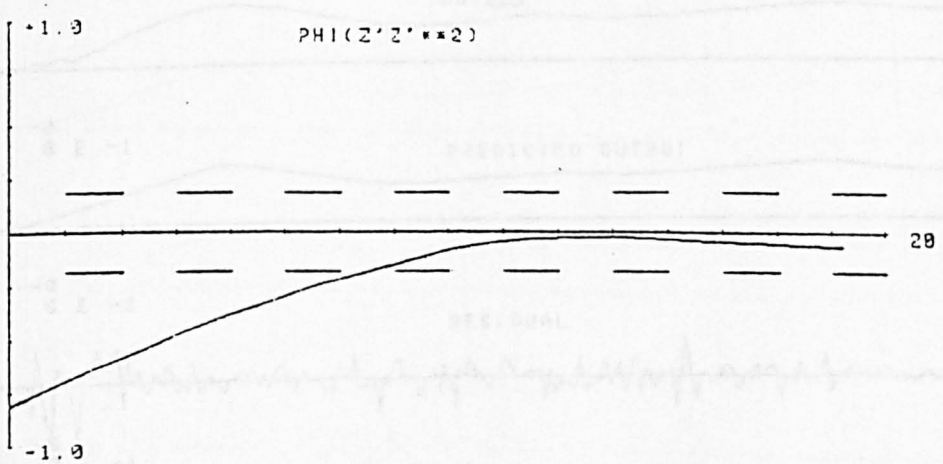


Figure 4.42: Structure detection test of the smoothed data

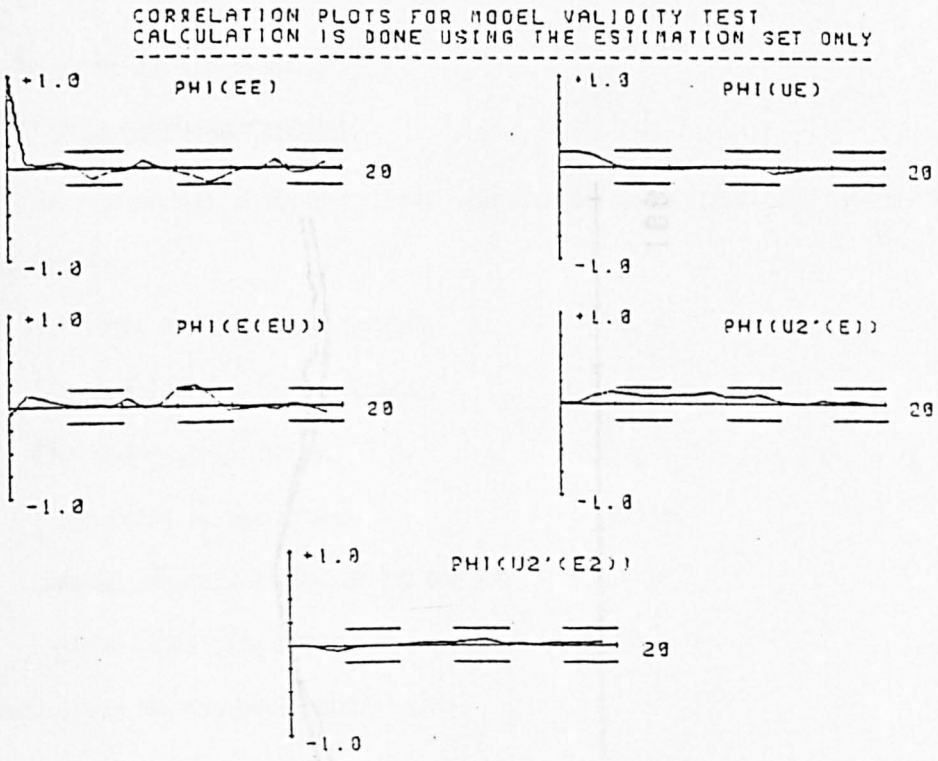


Figure 4.43: Model validity tests - best linear

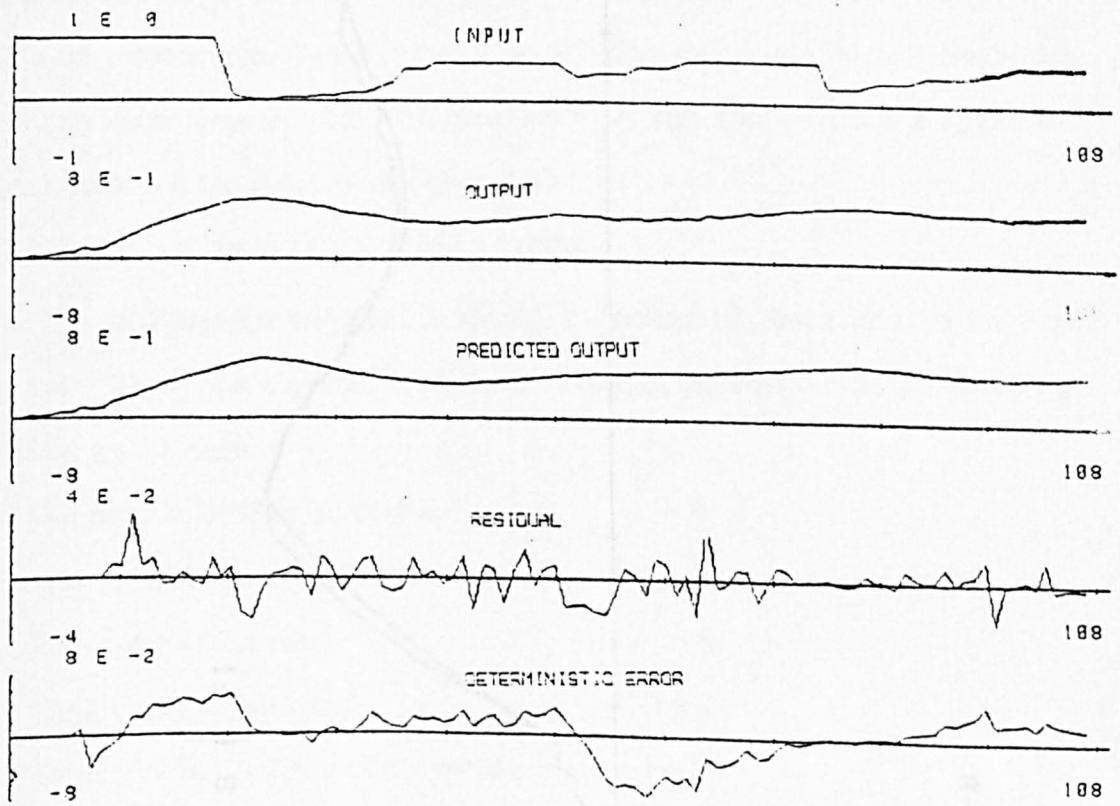


Figure 4.44: Plot of the residuals - best linear

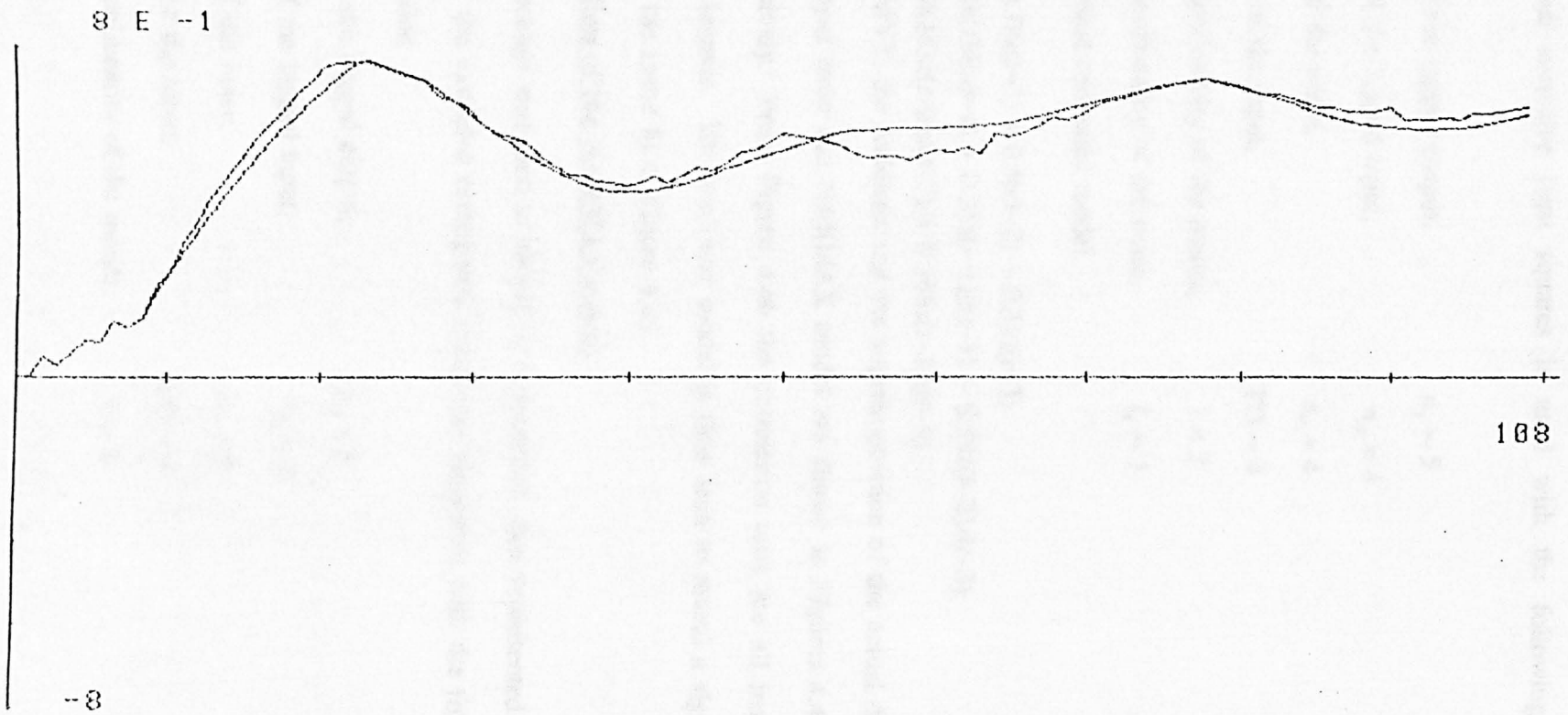


Figure 4.45: Process and predicted output superimposed - best linear

nonlinear terms are missing.

4.6.2) Best nonlinear model

An extended recursive least squares is used with the following initial specification

The order of the lagged output, $n_y = 5$

The order of the lagged input, $n_u = 4$

The order of the noise, $n_e = 4$

Time delay of the input, TD = 4

Degree of nonlinearity of the model, l = 2

Degree of nonlinearity of the noise, $l_e = 1$

which gives the final estimated model

$$\begin{aligned} z(t) = {} & 0.78z(t-1) + 0.5z(t-2) - 0.31z(t-3) \\ & + 0.034u(t-4) + 0.28z(t-1)u(t-7) - 0.39z(t-2)z(t-3) \\ & - 0.263z(t-2)u(t-7) + 0.366z(t-3)z(t-3) \end{aligned} \tag{4.9}$$

The plot of the MVT, the residuals and the superimposition of the actual data and the predicted output from this NARMAX model are shown in Figures 4.46, 4.47 and 4.48 respectively. From Figure 4.46 the correlation tests are all inside the 95% confidence bounds. The nonlinear model is thus seen to reveal a significant improvement on the linear fit of Figure 4.45.

4.6.3) Identification of the NARMAX model

The NOI package was used to identify the smoothed data represented in Figure 4.41. Using the extended orthogonal estimation algorithm with the following initinial specification

The order of the lagged output, $n_y = 4$

The order of the lagged input, $n_u = 4$

The order of the noise, $n_e = 5$

Time delay of the input, TD = 4

Degree of nonlinearity of the model, l = 2

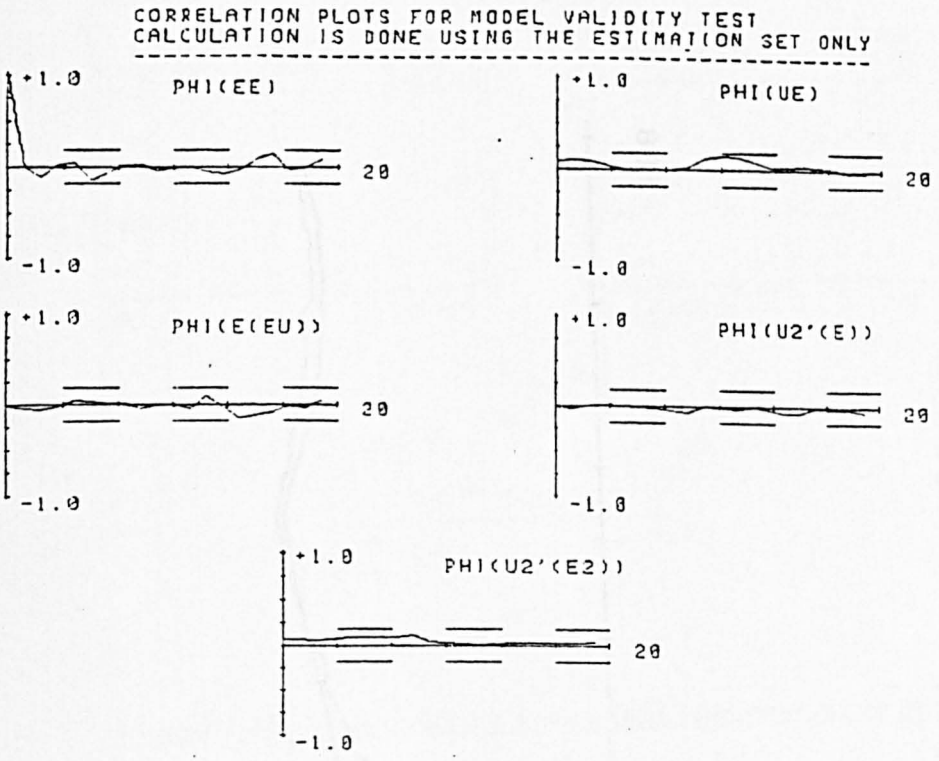


Figure 4.46: Model validity tests - best nonlinear

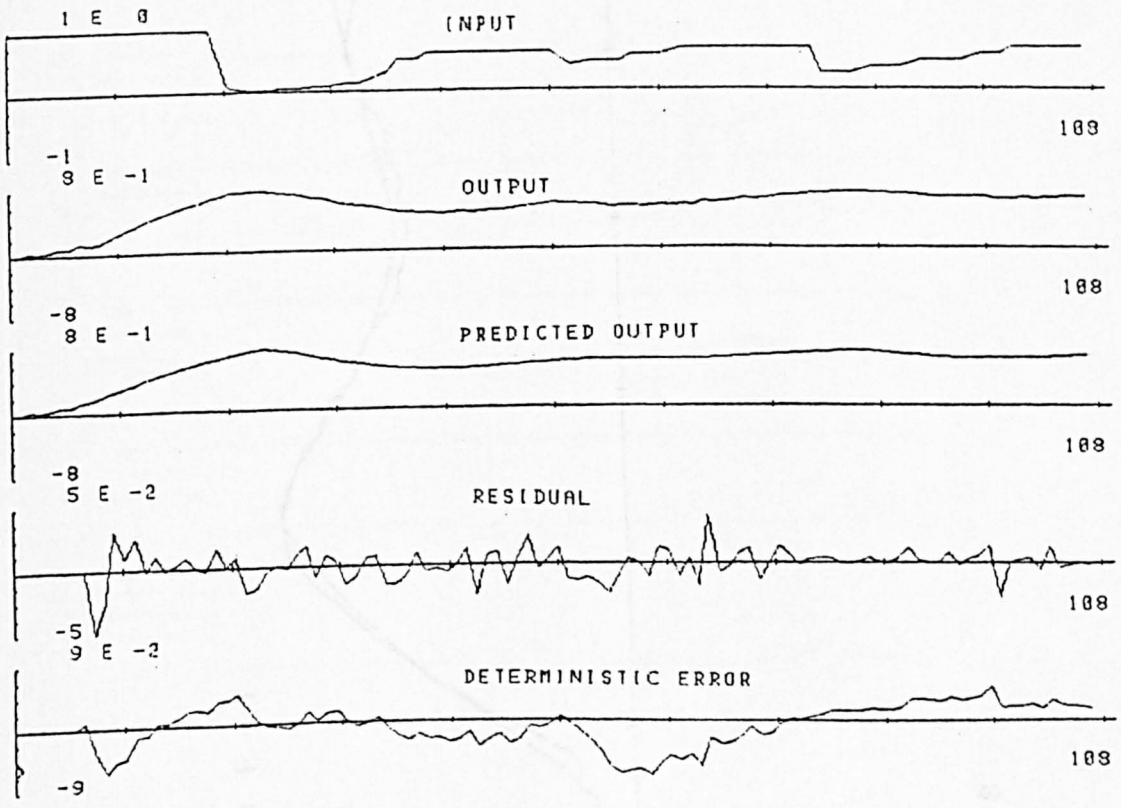


Figure 4.47: Plot of the residuals - best nonlinear

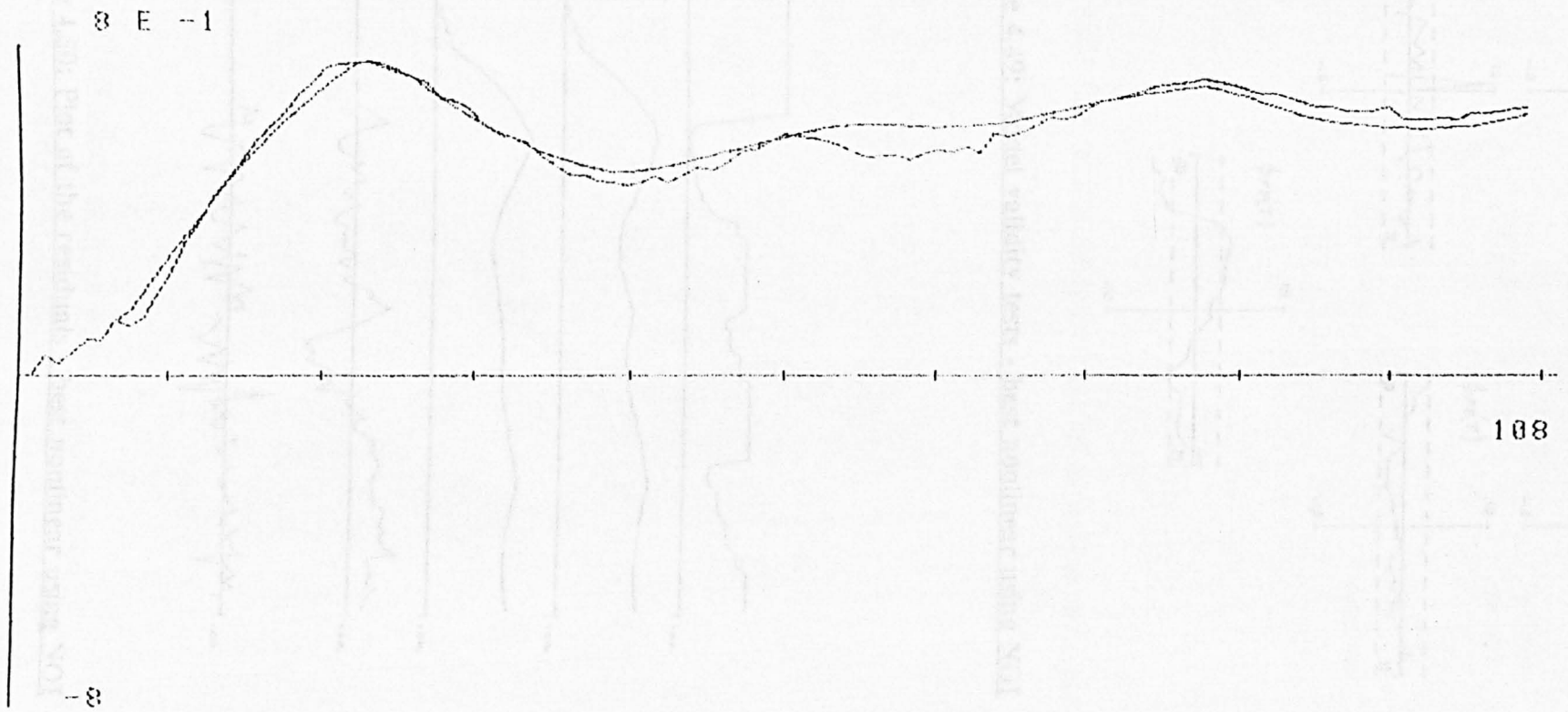


Figure 4.48: Process and predicted output superimposed - best nonlinear

PAGE MISSING IN ORIGINAL

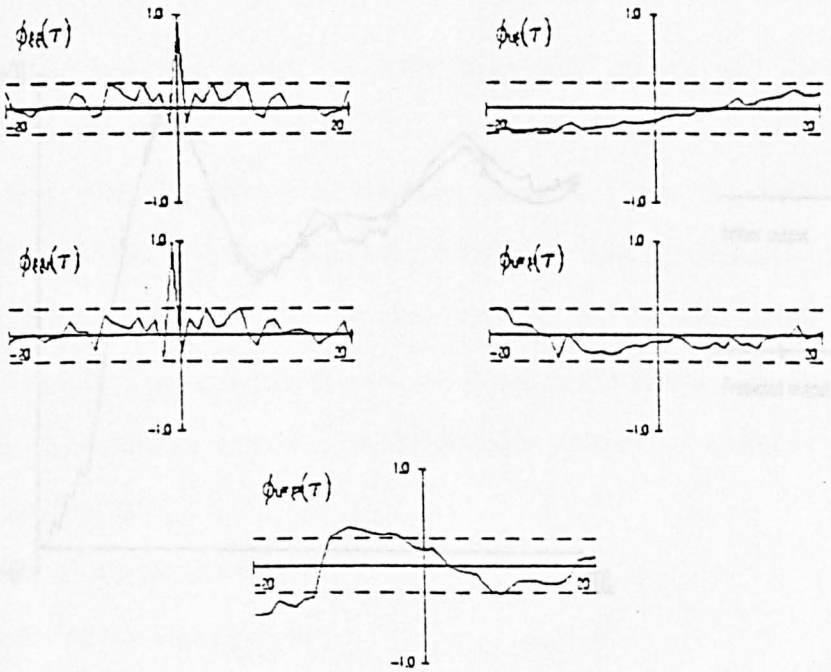


Figure 4.49: Model validity tests - best nonlinear using NOI

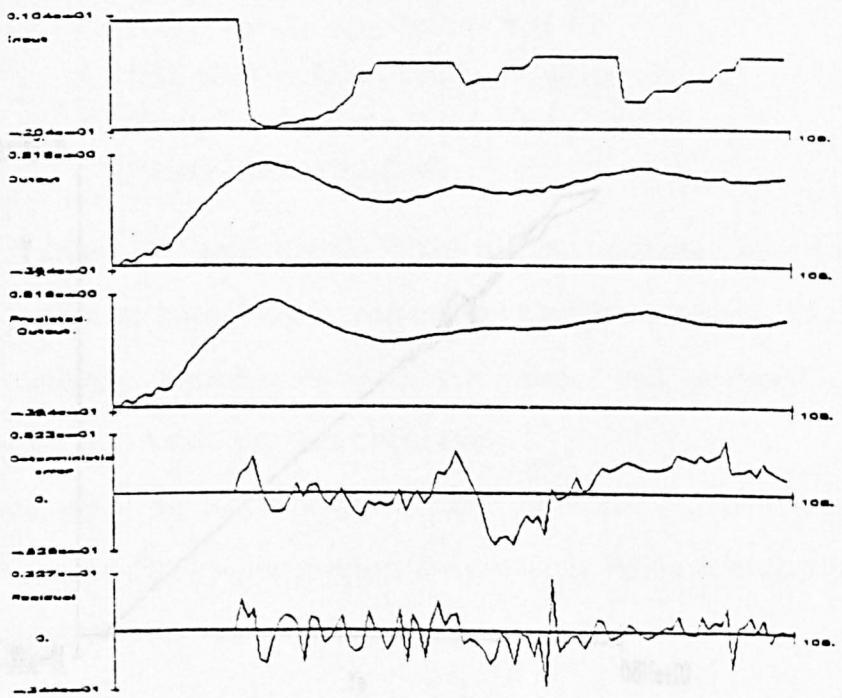


Figure 4.50: Plot of the residuals - best nonlinear using NOI

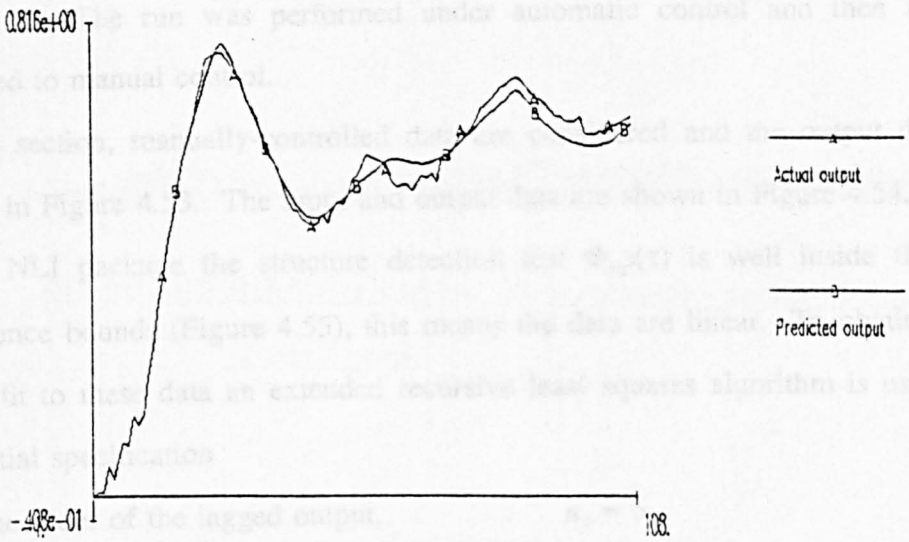


Figure 4.51: Process and predicted output superimposed - best nonlinear using NOI

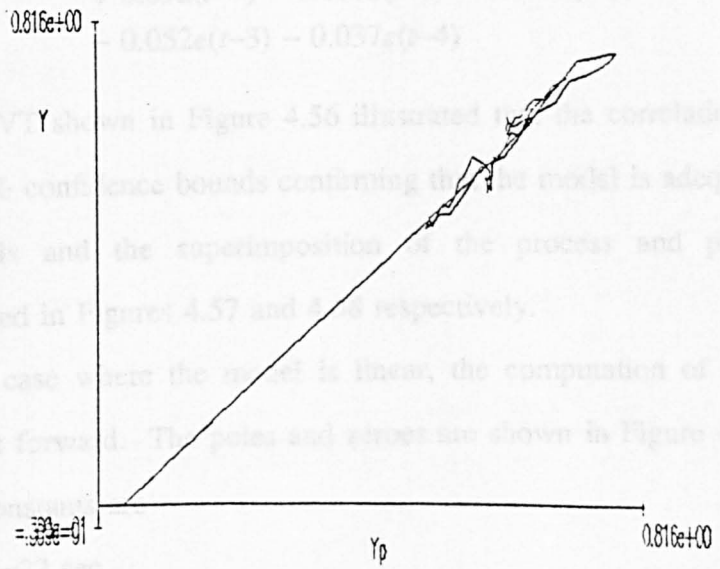


Figure 4.52: Relation between output and predicted output - best nonlinear

in 50 ml. The run was performed under automatic control and then initially switched to manual control.

In this section, manually-controlled data are considered and the output data are shown in Figure 4.53. The input and output data are shown in Figure 4.54.

Using NLI package the structure detection test $\Phi_{zz^2}(\tau)$ is well inside the 95% confidence bounds (Figure 4.55), this means the data are linear. To obtain a best linear fit to these data an extended recursive least squares algorithm is used with the initial specification

The order of the lagged output, $n_y = 6$

The order of the lagged input, $n_u = 4$

The order of the noise, $n_e = 4$

Time delay of the input, TD = 4

Degree of nonlinearity of the model, l = 1

which provides the following estimated model

$$\begin{aligned} z(t) = {} & 0.438z(t-1) + 0.3z(t-2) + 0.1z(t-3) \\ & - 0.048z(t-4) + 0.18z(t-5) - 0.09z(t-6) \\ & - 0.0005u(t-4) + 0.0097u(t-5) - 0.041u(t-6) \\ & + 0.09u(t-7) - 0.187e(t-1) - 0.104e(t-2) \\ & - 0.052e(t-3) - 0.037e(t-4) \end{aligned} \tag{4.12}$$

The MVT shown in Figure 4.56 illustrated that the correlation tests are all inside the 95% confidence bounds confirming that the model is adequate. The plot of the residuals and the superimposition of the process and predicted outputs are displayed in Figures 4.57 and 4.58 respectively.

In the case where the model is linear, the computation of the time constants is straight forward. The poles and zeroes are shown in Figure 4.59 and therefore the time constants are

τ_1 =22 sec

τ_2 =25 min.

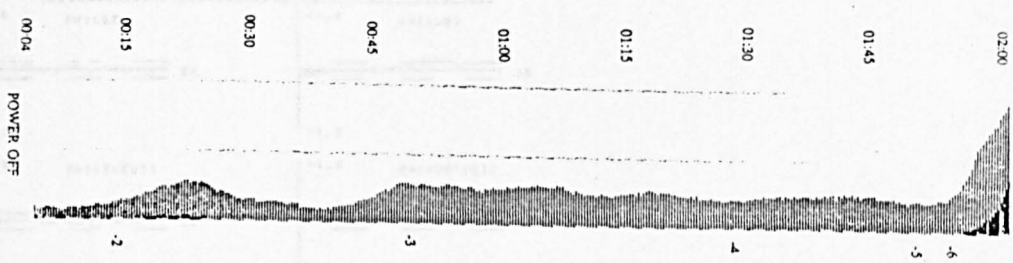


Figure 4.53: RIEMG recording from manually-controlled infusion of Atracurium

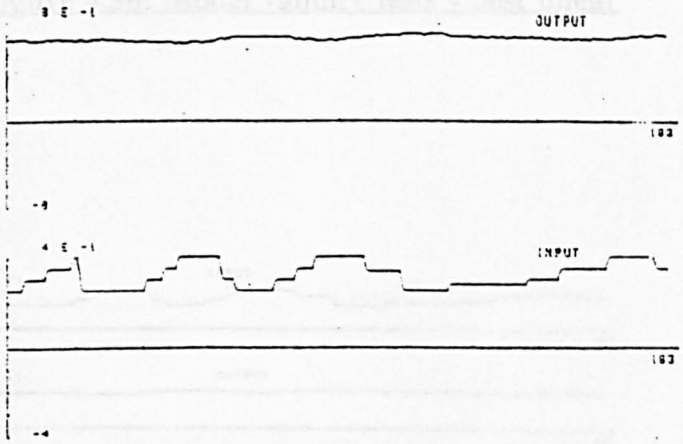


Figure 4.54: RIEMG and infusion data of Atracurium

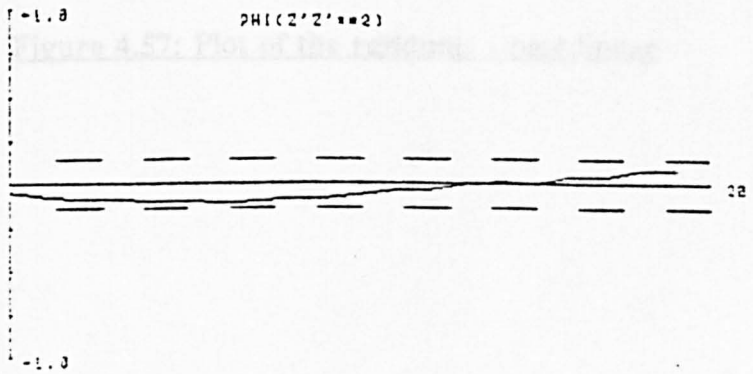


Figure 4.55: Structure detection test - best linear

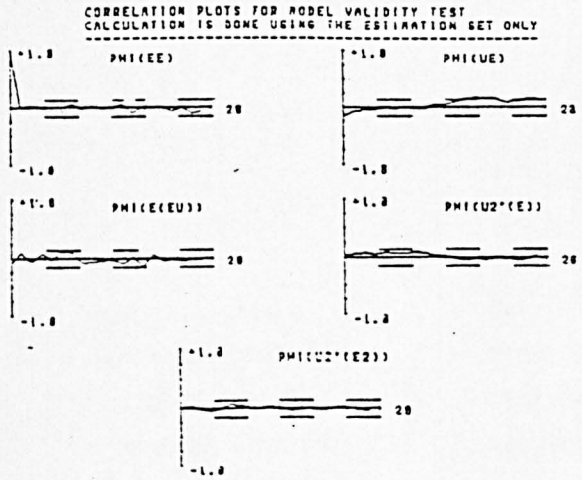


Figure 4.56: Model validity tests - best linear

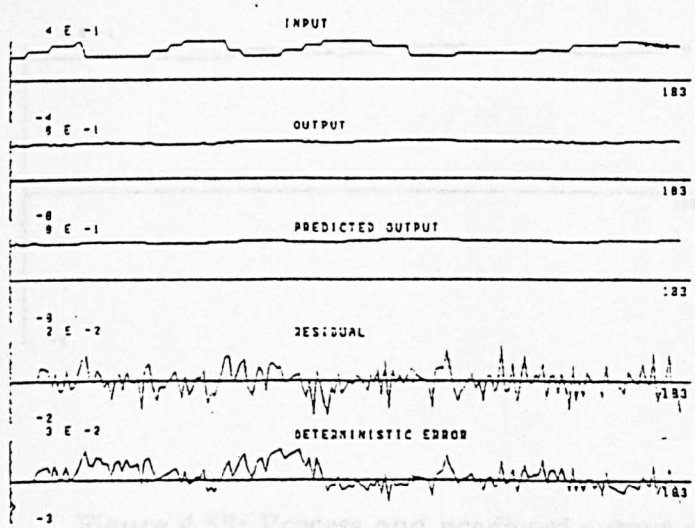


Figure 4.57: Plot of the residuals - best linear

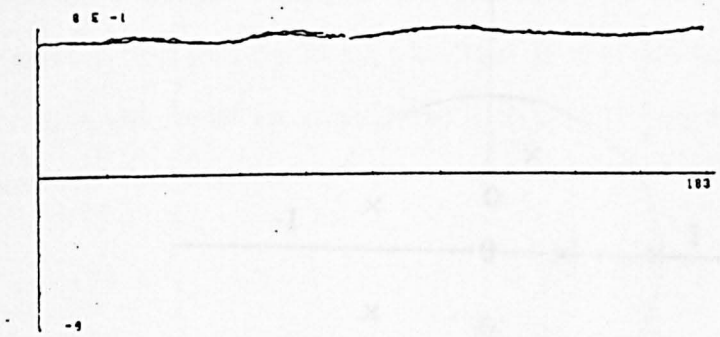


Figure 4.58: Process and predicted output superimposed - best linear

Poles		Zeros	
Real	*Imag*	*Real*	*Imag*
0.402E+00	0.000E+00	0.000E+00	0.000E+00
-0.622E+00	-0.404E+00	0.000E+00	0.000E+00
-0.622E+00	0.404E+00	0.211E-0.1	-0.316E+00
0.987E+00	0.000E+00	0.211E-0.1	0.316E+00
0.147E+00	-0.650E+00	0.196E-0.3	0.000E+00
0.147E+00	0.650E+00		

System Gain=2.461881

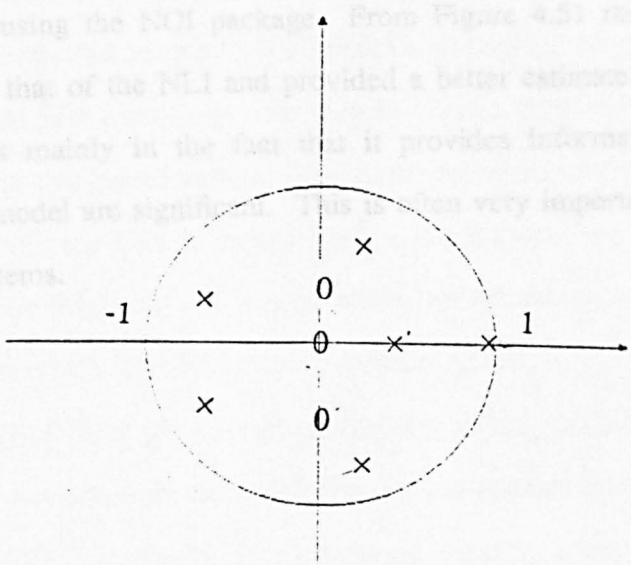


Figure 4.59: Poles and zeros of the linear model

4.8) Conclusion

Two extensive identification techniques, namely, NLI and NOI have been used to estimate a nonlinear difference equation (called NARMAX) model for the drugs Vecuronium and Atracurium.

For the case of Vecuronium, the NLI package was sufficient to use for identification of the model. The model validity tests shown in Figure 4.27 revealed that all the cross-correlation tests were inside the confidence bounds and the superimposition of the process and the predicted output (Figure 4.29) shows a very good fit.

For the case of Atracurium, the use of the NLI package offered a good model validity tests (Figure 4.43) and a fairly satisfactory fit (Figure 4.45). The fit was improved by using the NOI package. From Figure 4.51 the NOI has shown its superiority to that of the NLI and provided a better estimate. The strength of the algorithm lies mainly in the fact that it provides information regarding which terms in the model are significant. This is often very important in identification of nonlinear systems.

CHAPTER 5

Pharmacokinetics and pharmacodynamics identification for Vecuronium and Atracurium

5.1) Introduction

Although the NARMAX model gives a good fit to the data, comprising a small number of points, there is no clear relationship to the underlying physiological structure from this approach. It is widely considered that drug response models should comprise the linear pharmacokinetics to represent the drug distribution and absorption, and the nonlinear pharmacodynamics which are often modelled by a static characteristics using the Hill equation (Wagner, 1968, 1976). The overall characteristic (pharmacokinetics and pharmacodynamics) relating drug input to evoked EMG response was seen in chapter 2 to comprise a dead space to account for the margin of safety. These nonlinear effects may be embodied in a Hill equation (linear part in series with a dead space and saturation device) as shown in Figure 5.1, where α is a constant, not necessary an integer, and D is the value of X for which Y=50% of the maximum effect.

In this chapter, a fit of physiological structure to the models of both drugs, Vecuronium and Atracurium, is studied following the arrangement shown in Figure 5.2. Since the NARMAX model was estimated off-line, many data points could be used, and a white noise input with the mean adjusted to give the experimental data's RIEMG level.

An alternative approach to explore the underlying structure is to use cross-correlation methods based on Volterra series (Billing and Fakhouri, 1982). These methods require many data points (typically 5000 points) and Gaussian white noise inputs, which can be achieved, here, off-line using the identified NARMAX models. This technique should show whether the system under investigation is a

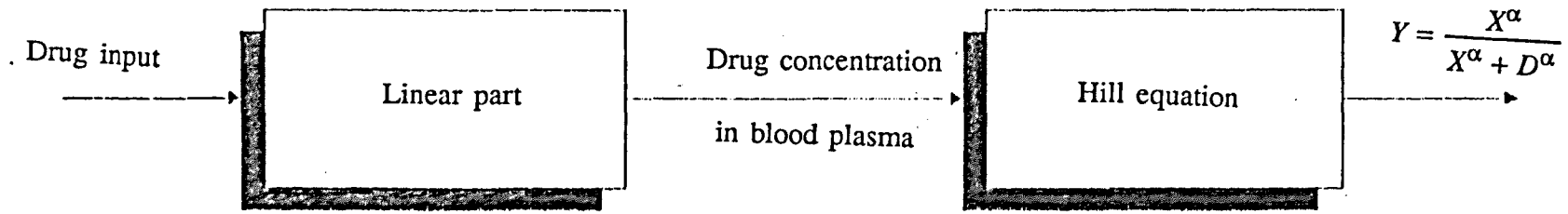


Figure 5.1: Representation of the overall characteristics

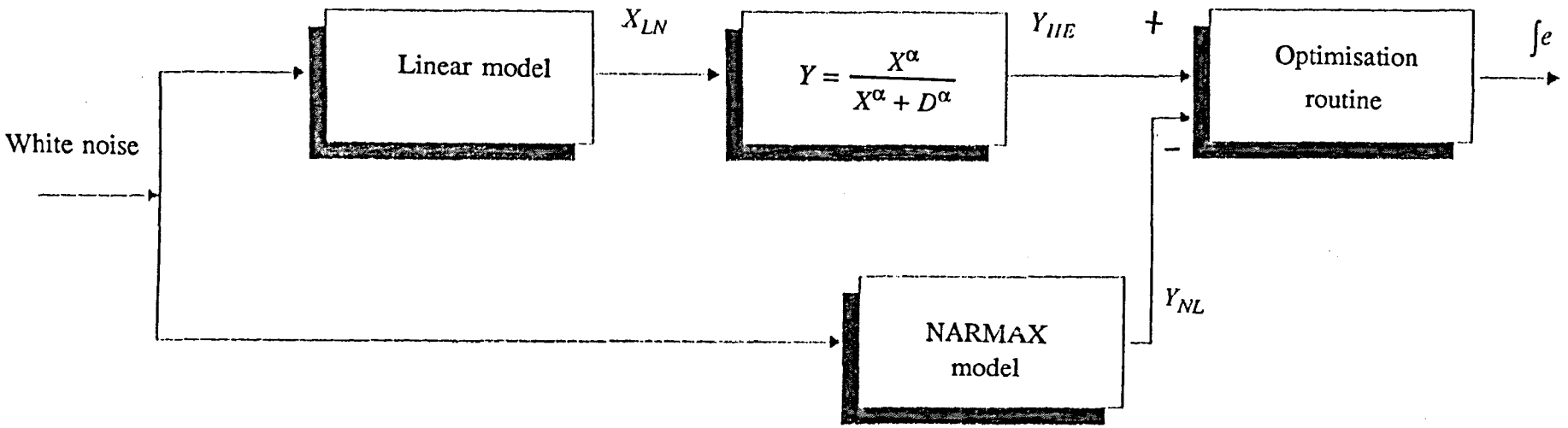


Figure 5.2: Schematic of methodology used for obtaining a physiologically meaningful model from the NARMAX model representation

Wiener or Hammerstein model. These two models will be derived in the next section.

5.2) Wiener and Hammerstein models

For the class of nonlinear systems such as cascade connections of linear dynamics and static nonlinear elements, system structure refers to the position of the nonlinear device in relation to the linear subsystems. Because the nonlinear Wiener (Billing and Fakhouri, 1977) and Hammerstein (Narendra and Gallman, 1966) models are all subclasses of the general models, an algorithm that can determine which, if any, of these models coincides with the structure of the process and provides an estimate of each component subsystem is desirable.

In this section, Weiner and Hammerstein structures are derived using a separable class of random processes method.

For the random variables X and Y let $f_1(x_1,\tau)$ and $f_2(y,\tau)$ be the probability density function and let $f(x,y,\tau)$ be the joint probability density functions. Let the conditional probability density function of X relative to Y in terms of the probability density functions of X and Y be

$$f_1(x/y;\tau) = \frac{f(x,y;\tau)}{f_2(y)} \tag{5.1}$$

$$f_2(y/x;\tau) = \frac{f(x,y;\tau)}{f_1(x)} \tag{5.2}$$

hence

$$\begin{aligned} f(x,y;\tau) &= f_1(x/y;\tau)\, f_2(y) \\ &= f_2(y/x;\tau)\, f_1(x) \end{aligned} \tag{5.3}$$

If Y is a single-value function of the random variable X and is given by

$$y = \Psi(x) \tag{5.4}$$

then for any fixed value x of the variable X , Y possesses a single possible value $\Psi(x)$ and the conditional probability of this value is equal to unity.

Consequently, the conditional probability density function of the random variable Y with respect to X is the delta (δ) function given by

$$f_2(y/x;\tau) = \delta(y-\Psi(x)) \tag{5.5}$$

$$f_2(y/x;\tau) = \begin{cases} 1 & \text{if } y = \Psi(x) \\ 0 & \text{if } y \neq \Psi(x) \end{cases}$$

using equation 5.3

$$f(x,y;\tau) = f_1(x)\delta(y-\Psi(x)) \tag{5.6}$$

(Pugachev, 1962)

Also define

$$g(y;\tau) = \int_{-\infty}^{\infty} xf(x,y;\tau)dx \tag{5.7}$$

x(t) is separable with respect to y(t) if

$$g(y;\tau) = g_1(y)g_2(\tau) \tag{5.8}$$

using equation 5.8 and 5.6

$$g(y,\tau) = \int_{-\infty}^{\infty} xf(x,y;\tau)dx$$

$$g(y,\tau) = \int_{-\infty}^{\infty} xf_1(x)\delta(y-\Psi(x))dx \tag{5.9}$$

If Ψ is a one to one function with Ψ^{-1} its inverse

$$g(y,\tau) = \Psi^{-1}(y)f_1(\Psi^{-1}(y)) \tag{5.10}$$

for τ =0 equations 5.8 and 5.10 become

$$g(y,0) = g_1(y)g_2(0) \tag{5.11}$$

$$g(y,0) = \Psi^{-1}(y)f_1 o \Psi^{-1}(y) \tag{5.12}$$

hence

$$g_1(y) = \frac{\Psi^{-1}(y)f_1 o \Psi^{-1}(y)}{g_2(0)} \tag{5.13}$$

Now consider the cross-correlation function

$$\Phi_{xy}(\tau) = E[X(t)Y(t+\tau)]$$

$$\Phi_{xy}(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y;\tau)dxdy$$

$$\Phi_{xy}(\tau) = \int_{-\infty}^{\infty} y\left(\int_{-\infty}^{\infty} xf(x,y;\tau)dx\right)dy$$

$$\Phi_{xy}(\tau) = \int_{-\infty}^{\infty} yg(y,\tau)dy$$

$$\Phi_{xy}(\tau) = \int_{-\infty}^{\infty} yg_1(y)g_2(\tau)dy \tag{5.14}$$

Since the process is separable, using equation 5.8

$$\Phi_{xy}(\tau) = g_2(\tau)\int_{-\infty}^{\infty} yg(y)dy \tag{5.15}$$

For τ =0 equation 5.15 becomes

$$\Phi_{xy}(0) = g_2(0)\int_{-\infty}^{\infty} yg(y)dy \tag{5.16}$$

Hence

$$g_2(\tau) = \frac{g_2(0)\Phi_{xy}(\tau)}{\Phi_{xy}(0)} \tag{5.17}$$

The idea can be illustrated by considering the general model of Figure 3.2 of section 3.2.2.

Assume that x(t) can be measured and let $f(x_1,x_2;\tau)$ be the second order probability density function of the process x(t) and analogous to equations 5.7 and 5.8

$$g(x_2,\tau) = \int_{-\infty}^{\infty} x_1 f(x_1,x_2;\tau)dx_1, \text{ for } all\ x,\tau \tag{5.18}$$

$$g(x_2,\tau) = g_1(x_2)g_2(\tau), \text{ for } all\ x,\tau$$

and analogous to equations 5.13 and 5.17

$$g_1(x_2) = \frac{x_2 f(x_2)}{g_2(0)} \tag{5.19}$$

$$g_2(\tau) = \frac{g_2(0)\Phi_{x_1x_2}(\tau)}{\Phi_{x_1x_2}(0)} \tag{5.20}$$

Define the cross-correlation function

$$\Phi_{xq}(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 F[x_2]\, f(x_1,x_2,\tau)dx_1dx_2 \tag{5.21}$$

Using equations 5.18, 5.19 and 5.20

$$\Phi_{xq}(\tau) = \int_{-\infty}^{\infty} F[x_2]g_1(x_2)g_2(\tau)dx_2$$

$$\Phi_{xq}(\tau) = \frac{\Phi_{x_1x_2}(\tau)}{\Phi_{x_1x_2}(0)} \int_{-\infty}^{\infty} F[x_2]x_2 f(x_2)dx_2$$

$$\Phi_{xq}(\tau) = C_F \Phi_{x_1x_2}(\tau) \tag{5.22}$$

where

C_F is a constant.

This is known as the invariance property (Billings and Fakhouri, 1978)

But in practice the internal signals in Figure 3.2 x(t) and q(t) will not be available for measurement. However by considering separability under linear and nonlinear transformations equation 5.22 can be generalised. For a non-zero mean Gaussian white noise input Figure 3.2 yields

$$\Phi_{uy}(\tau) = C_{FG} \int_{-\infty}^{\infty} h_1(\tau_1)h_2(\tau-\tau_1)d\tau_1 \tag{5.23}$$

$$\Phi_{u^2y}(\tau) = C_{FFG} \int_{-\infty}^{\infty} h_1^2(\tau-\tau_1)h_2(\tau_1)d\tau_1 \tag{5.24}$$

where

C_{FG} and C_{FFG} are constants, providing $h_1(t)$ are stable bounded-inputs bounded-outputs.

Equation 5.23 represents the first degree correlation function and equation 5.24 represents the second degree correlation function. If the second degree correlation function is the square of the first degree correlation function , except for a constant of proportionality (setting $h_2(t) = \delta(t)$ in equations 5.23 and 5.24), the structure refers to a Weiner model (linear dynamics followed by a nonlinear static characteristic).

However, if the first and second degree correlation functions are equal except for a constant of proportionality (setting $h_1(t) = \delta(t)$ in equations 5.23 and 5.24), the structure is consistant with a Hammerstein model (nonlinear static characteristic followed by linear dynamics) (Billings, 1985).

5.3) Fitting a physiological structure to the Vecuronium model

It has been mentioned in section 5.1 that although the NARMAX model provided a good fit to the data comprising a small number of points, it by no means reflects the under lying physiological structure. To obtain parameters for this structure, the arrangement shown in Figure 5.2 is used.

The pharmacokinetics were modelled in two ways. First, a linear model was obtained by truncating the first few points and represented by equation 4.4, and second, the model was obtained by linearization of the NARMAX model about an operating point.

5.3.1) Use of the linear model obtained by truncating data

Consider the 'best' nonlinear model (equation 4.5) and linear model (equation 4.4). A white noise sequence is used as input for both models (Figure 5.3).

To find the parameters of the Hill equation, an optimization method, the simplex method, available in the NAG library was used.

Prior to optimization, Y_{NL} was plotted as a function of X_{LN}

where

Y_{NL} is the output of the nonlinear model

X_{LN} is the output of the linear model

Y_{HE} is the output of the Hill equation block

to give the results shown in Figures 5.4 and 5.5 respectively. The nonlinear static response (related to the pharmacodynamics) could now be estimated by fitting a Hill equation.

After optimization, the parameters of the Hill equation turned out to be

$$\alpha = 0.8629 \quad , \quad D = 0.4455$$

By substituting the values of α and D into the Hill equation, Y_{HE} and Y_{NL} are shown in Figure 5.6 and the phase plot Y_{HE}/X_{LN} is shown in Figures 5.7 and 5.8 respectively.

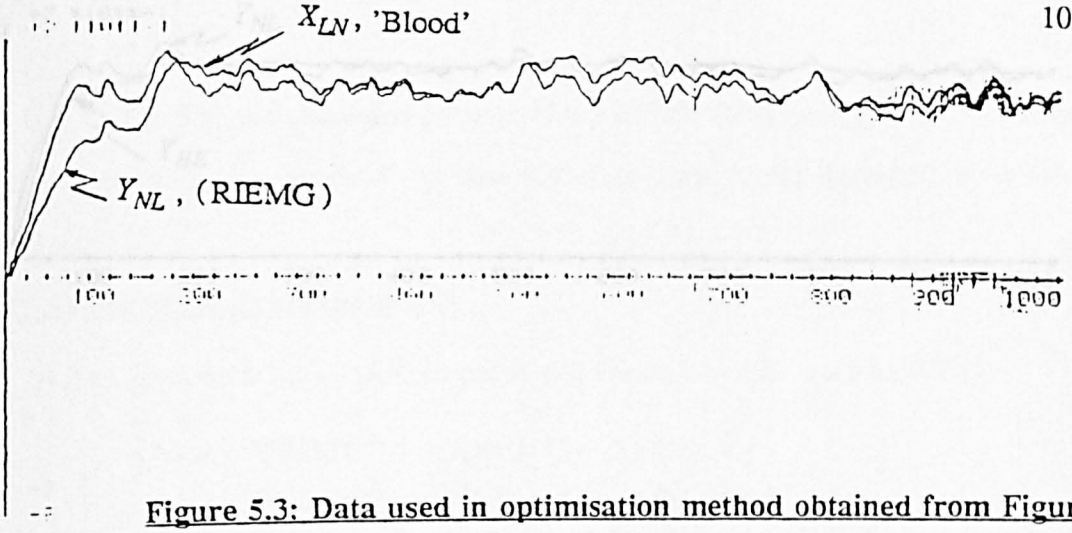


Figure 5.3: Data used in optimisation method obtained from Figure 5.2

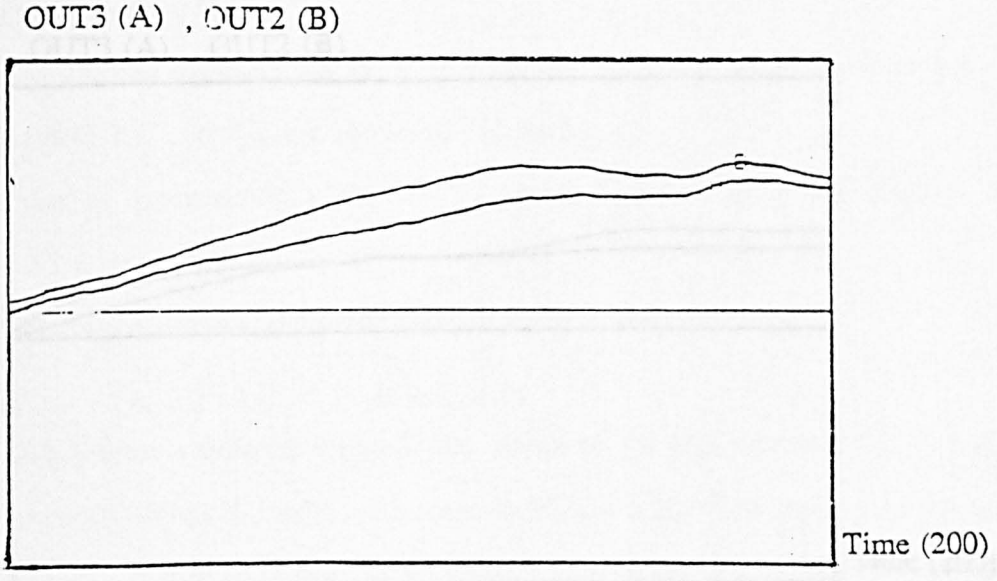


Figure 5.4: Representation of Y_{NL} and X_{LN}

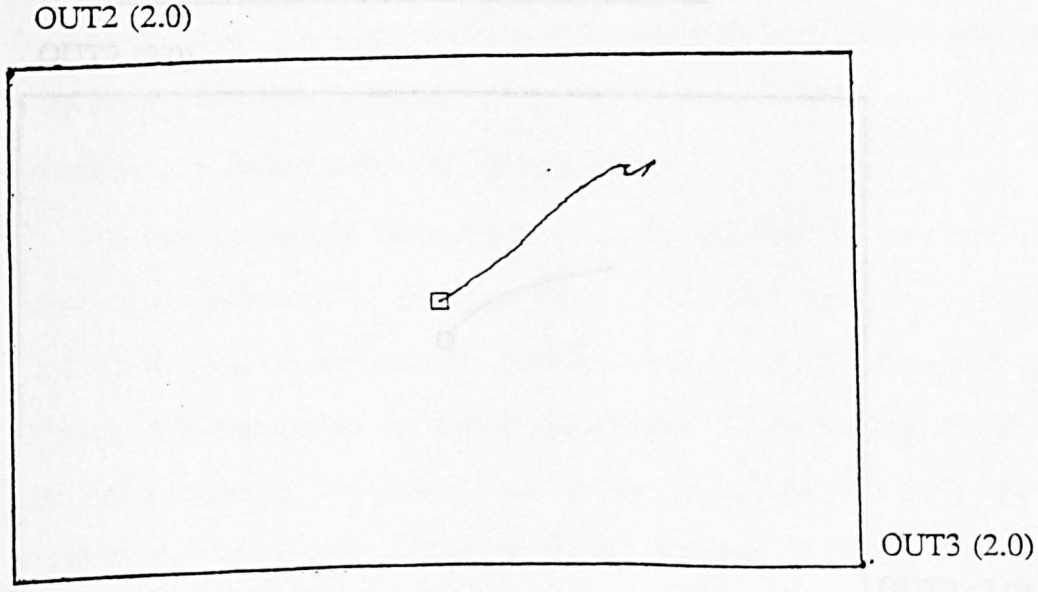


Figure 5.5: Phase plot of Y_{NL} versus X_{LN}

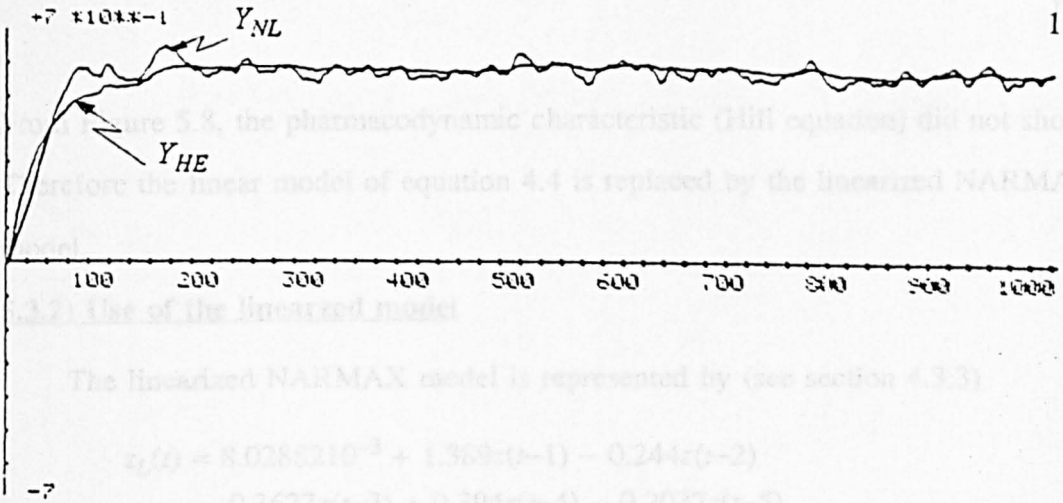


Figure 5.6: Data obtained after optimisation

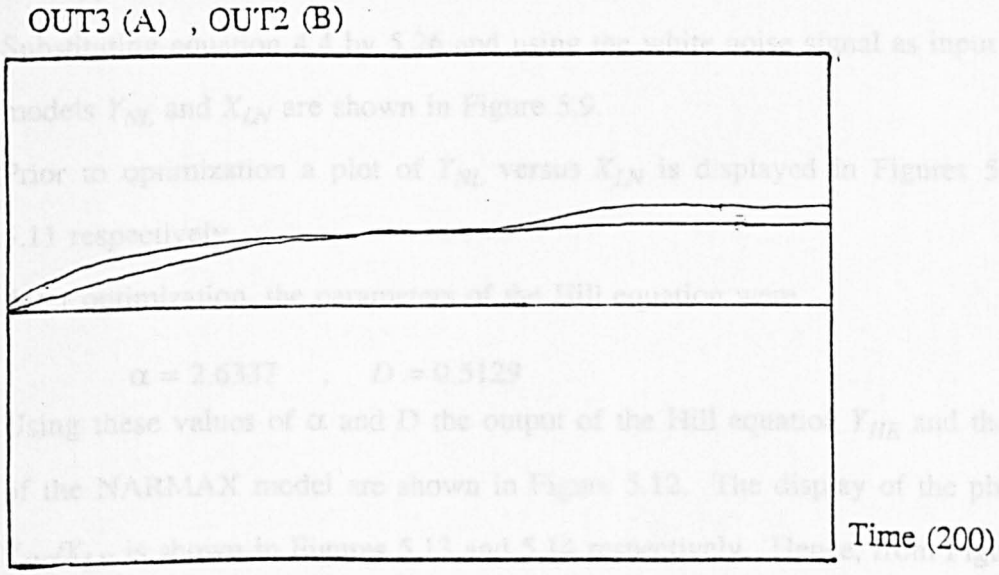


Figure 5.7: Representation of Y_{HE} and X_{LN}

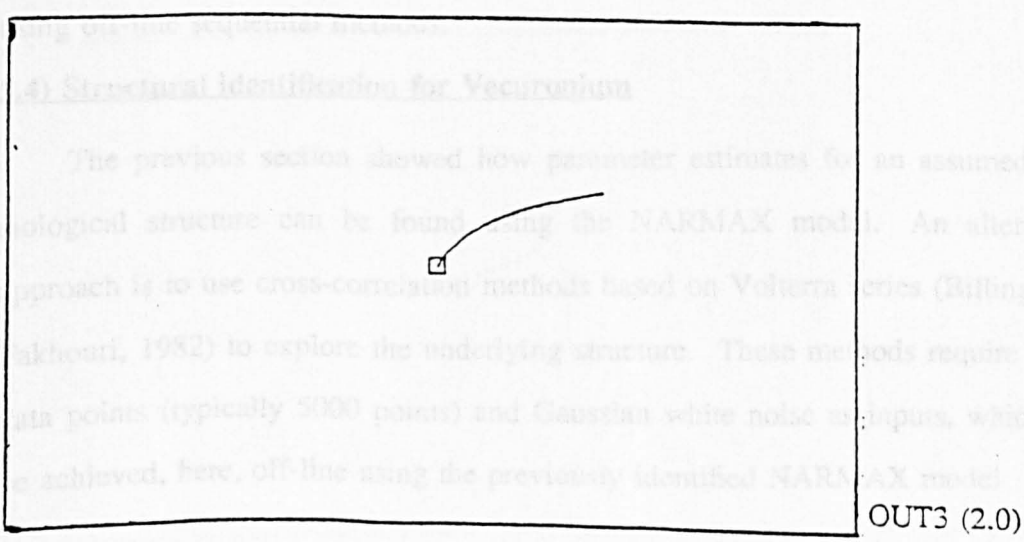


Figure 5.8: Phase plot of Y_{HE} versus X_{LN}

From Figure 5.8, the pharmacodynamic characteristic (Hill equation) did not show. Therefore the linear model of equation 4.4 is replaced by the linearized NARMAX model.

5.3.2) Use of the linearzed model

The linearized NARMAX model is represented by (see section 4.3.3)

$$\begin{aligned} z_L(t) = {} & 8.0288210^{-3} + 1.389z(t-1) - 0.244z(t-2) \\ & - 0.3627z(t-3) + 0.394z(t-4) - 0.2037z(t-5) \\ & + 8.11310^{-3}u(t-3) + 4.778810^{-3}u(t-4) + 2.17510^{-3}u(t-5) \end{aligned} \qquad (5.25)$$

(see appendix C).

Substituting equation 4.4 by 5.26 and using the white noise signal as input to both models Y_{NL} and X_{LN} are shown in Figure 5.9.

Prior to optimization a plot of Y_{NL} versus X_{LN} is displayed in Figures 5.10 and 5.11 respectively.

After optimization, the parameters of the Hill equation were

$$\alpha = 2.6337 \quad , \quad D = 0.5129$$

Using these values of α and D the output of the Hill equation Y_{HE} and the output of the NARMAX model are shown in Figure 5.12. The display of the phase plot Y_{HE}/X_{LN} is shown in Figures 5.13 and 5.14 respectively. Hence, from Figure 5.14, the phase plot shows the characteristic of the Hill equation. The values of α and D are respectable and commensurate with values obtained in previous studies using off-line sequential methods.

5.4) Structural identification for Vecuronium

The previous section showed how parameter estimates for an assumed physiological structure can be found using the NARMAX model. An alternative approach is to use cross-correlation methods based on Volterra series (Billings and Fakhouri, 1982) to explore the underlying structure. These methods require many data points (typically 5000 points) and Gaussian white noise as inputs, which can be achieved, here, off-line using the previously identified NARMAX model

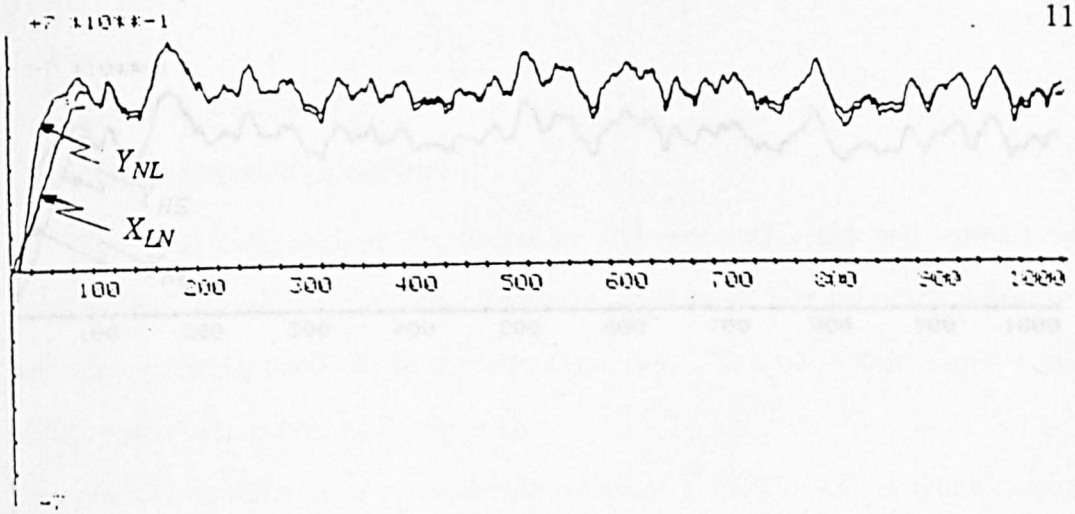


Figure 5.9: Data used in optimisation method

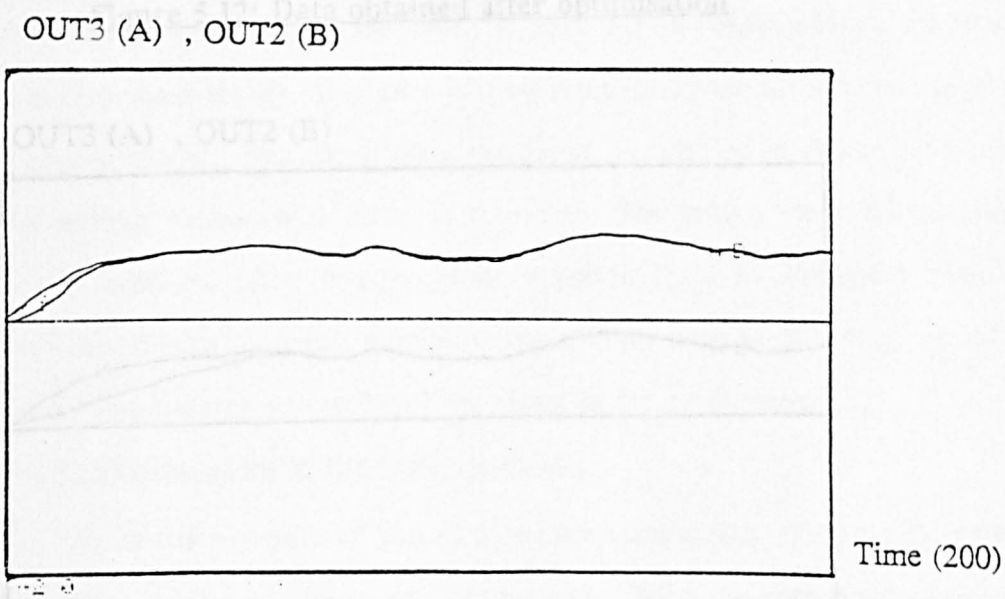


Figure 5.10: Representation of Y_{HE} and X_{LN}

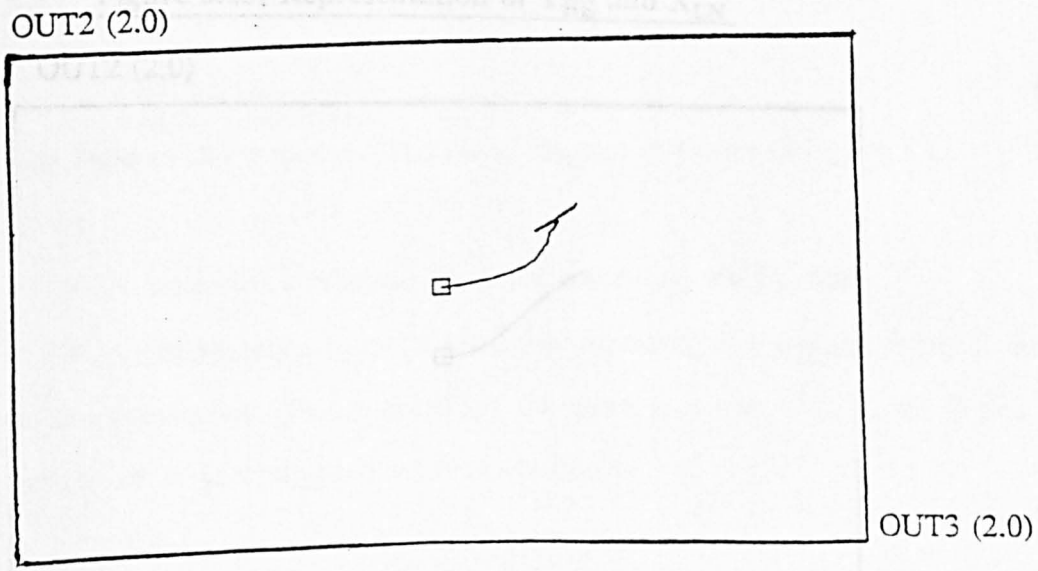


Figure 5.11: Phase plot of Y_{NL} versus X_{LN}

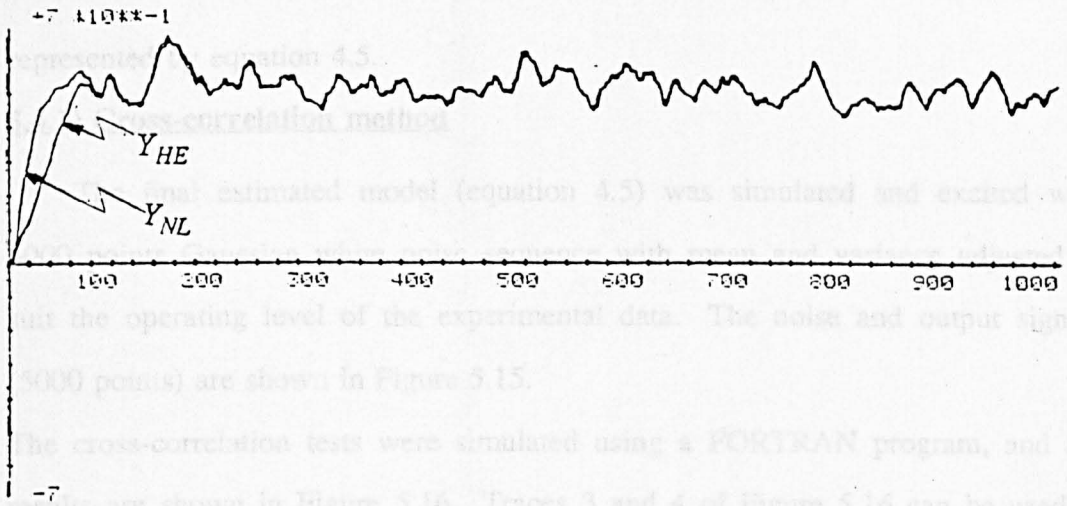


Figure 5.12: Data obtained after optimisation

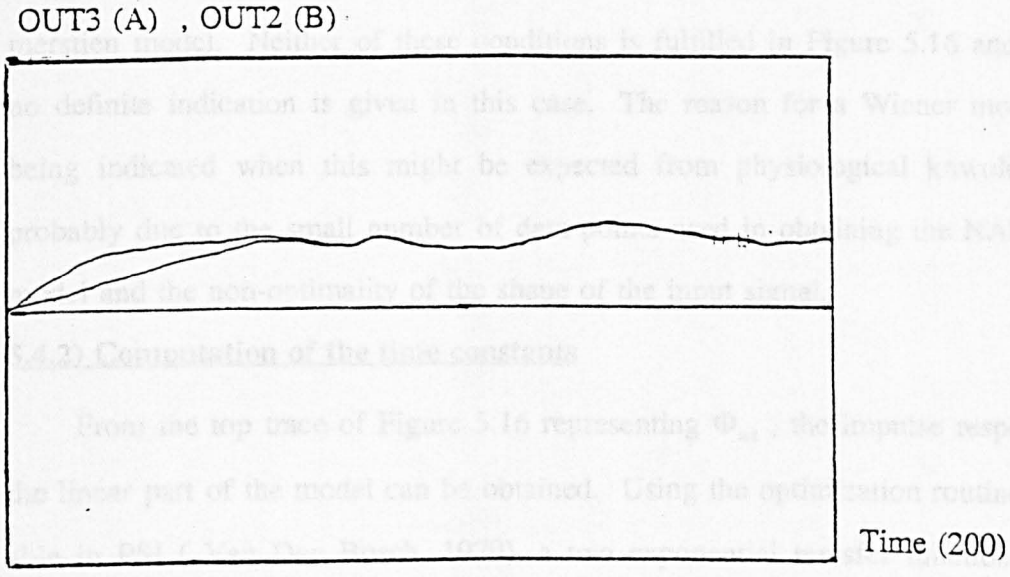


Figure 5.13: Representation of Y_{HE} and X_{LN}

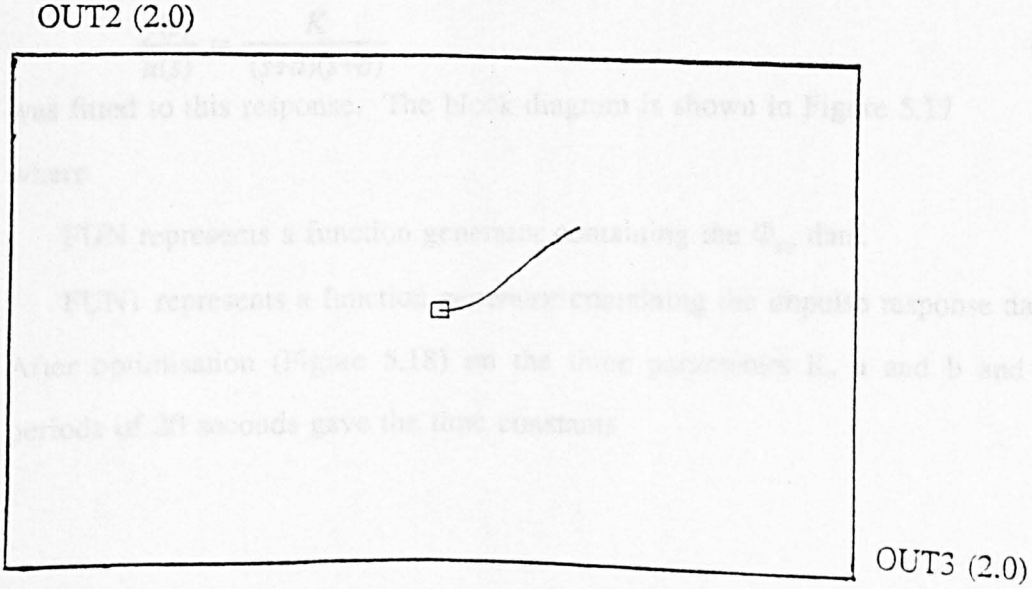


Figure 5.14: Phase plot of Y_{HE} versus X_{LN}

represented by equation 4.5.

5.4.1) Cross-correlation method

The final estimated model (equation 4.5) was simulated and excited with 5000 points Gaussian white noise sequence with mean and variance adjusted to suit the operating level of the experimental data. The noise and output signals (5000 points) are shown in Figure 5.15.

The cross-correlation tests were simulated using a FORTRAN program, and the results are shown in Figure 5.16. Traces 3 and 4 of Figure 5.16 can be used to determine the underlying structure. If trace 3 is a constant level, the structure is that of Wiener model. If trace 4 is a constant level then the structure is of a Hammerstien model. Neither of these conditions is fulfilled in Figure 5.16 and hence no definite indication is given in this case. The reason for a Wiener model not being indicated when this might be expected from physiological knwoledge is probably due to the small number of data points used in obtaining the NARMAX model and the non-optimality of the shape of the input signal.

5.4.2) Computation of the time constants

From the top trace of Figure 5.16 representing Φ_{uz} , the impulse response of the linear part of the model can be obtained. Using the optimization routine available in PSI (Van Den Bosch, 1979), a two exponential transfer function of the form

$$\frac{y(s)}{u(s)} = \frac{K}{(s+a)(s+b)} \tag{5.26}$$

was fitted to this response. The block diagram is shown in Figure 5.17

where

FUN represents a function generator containing the Φ_{uz} data.

FUN1 represents a function generator containing the impulse response data.

After optimisation (Figure 5.18) on the three parameters K, a and b and using periods of 20 seconds gave the time constants

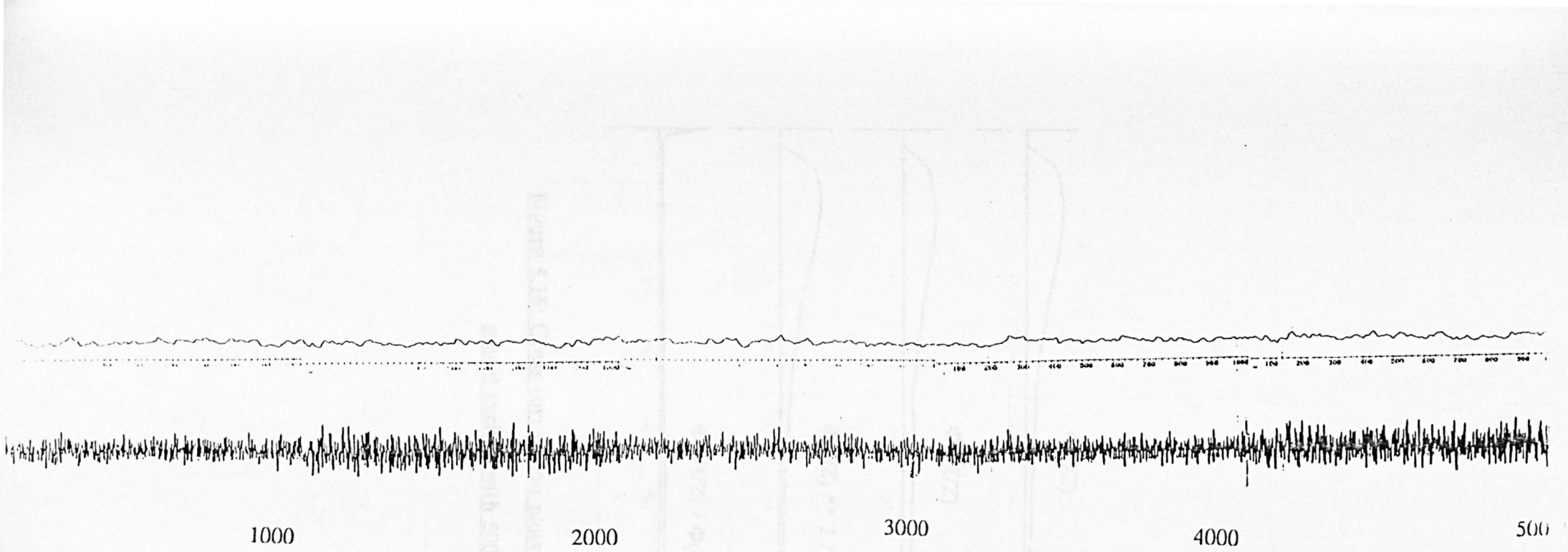


Figure 5.15: Representation of the 5000 points white noise sequence

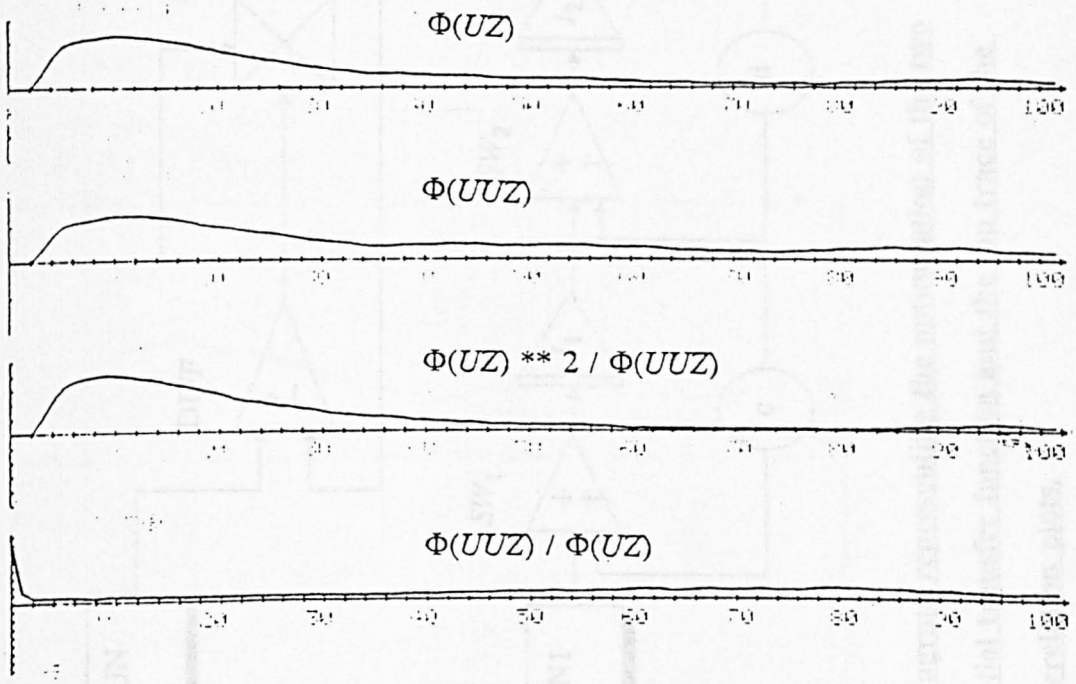


Figure 5.16: Cross correlation plots obtained from the NARMAX model excited with 5000 point white noise sequence

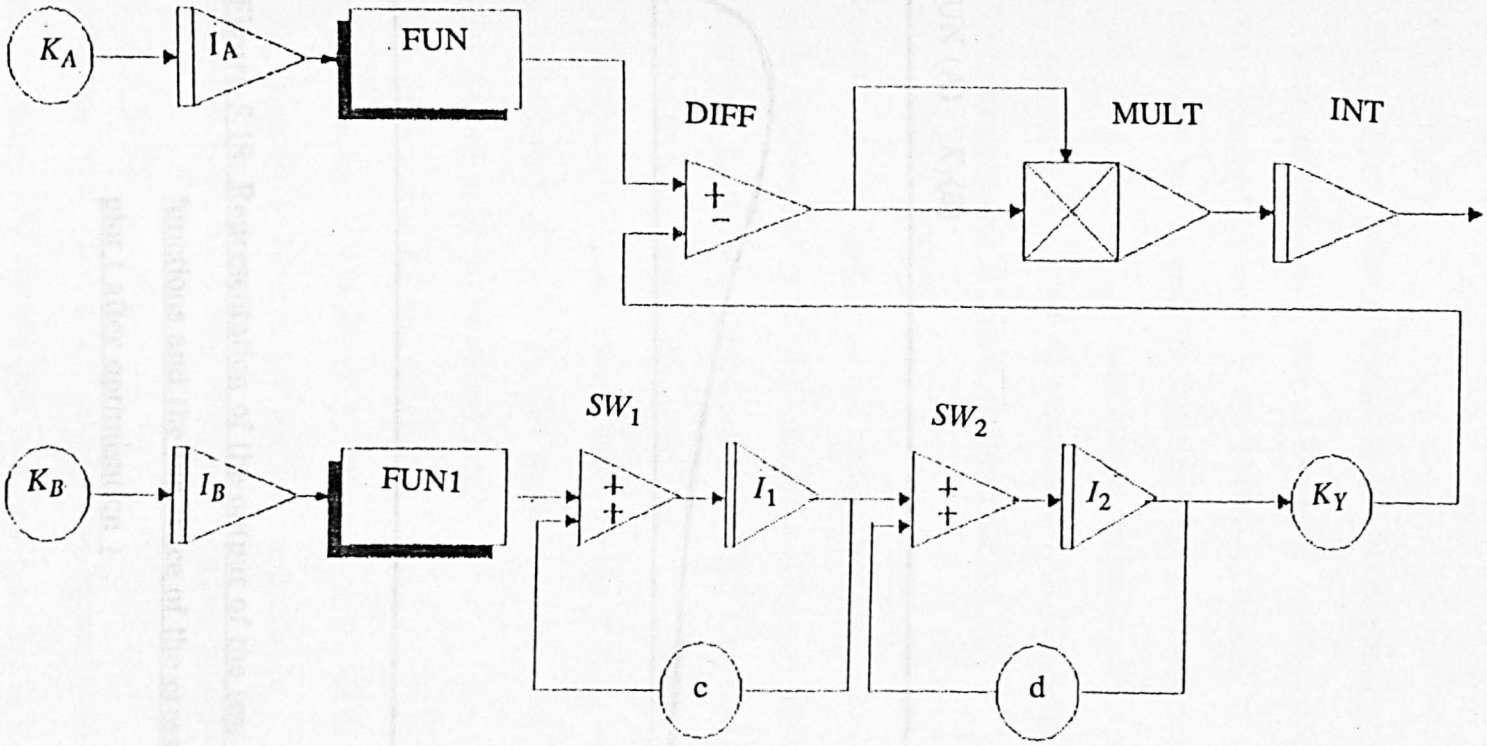


Figure 5.17: Block diagram representing the optimisation of the two exponential transfer function and the top trace of the cross-correlation plots.

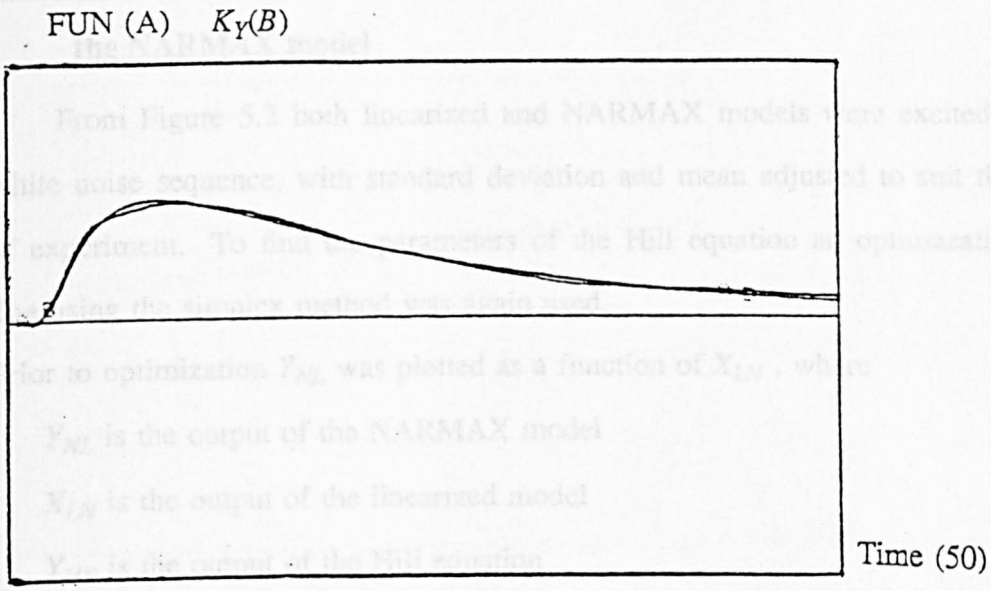


Figure 5.18: Representation of the output of the two exponential transfer functions and the top trace of the cross-correlation plot (after optimisation)

$\tau_1 = 0.96$ min , $\tau_2 = 6.31$ min

These values are comensurate with pharmacokinetics of Vecuronium obtained via blood assay methods.

5.5) Fitting a physiological structure to the Atracurium model (NLI)

The same method is used here as in section 5.3 for Vecuronium . Many data points were obtained by exciting the NARMAX model of equation 4.7 on the whole data (218 points) and the linearized model of equation 4.8. The arrangement is shown in Figure 5.2. Both models are excited by a white noise sequence with its mean adjusted to give the experimental RIEMG level.

5.5.1) The linear model obtained by linearization of the NARMAX model

From Figure 5.2 both linearized and NARMAX models were excited with a white noise sequence, with standard deviation and mean adjusted to suit the level of experiment. To find the parameters of the Hill equation an optimazation routine using the simplex method was again used.

Prior to optimization Y_{NL} was plotted as a function of X_{LN} , where

Y_{NL} is the output of the NARMAX model

X_{LN} is the output of the linearized model

Y_{HE} is the output of the Hill equation

Before optimization, the display of Y_{NL} and X_{LN} is shown in Figure 5.19 and Y_{NL} versus X_{LN} is shown in Figures 5.20 and 5.21 respectively.

After optimization the parameters of the Hill equation turned out to be

$\alpha = 2.131$, $D = 0.405$

Substituting the values of α and D into the Hill equation, the plot of Y_{HE} and Y_{NL} is shown in Figure 5.22 and the plot of Y_{HE} versus X_{LN} is shown in Figures 5.23 and 5.24 respectively.

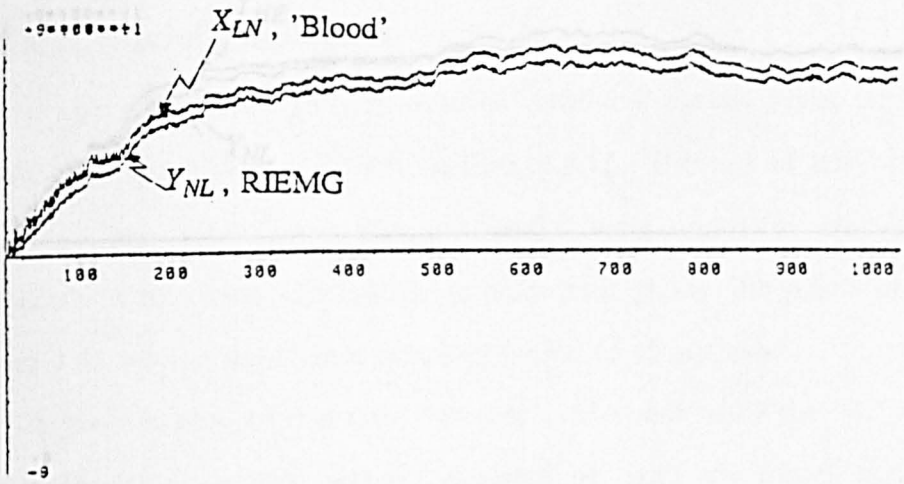


Figure 5.19: Data used in optimisation method obtained from Figure 5.2

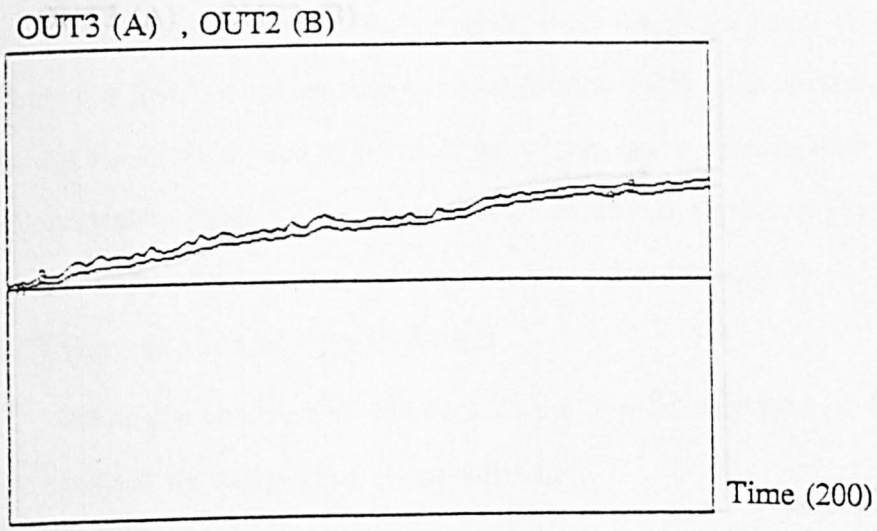


Figure 5.20: Representation of Y_{NL} and X_{LN}

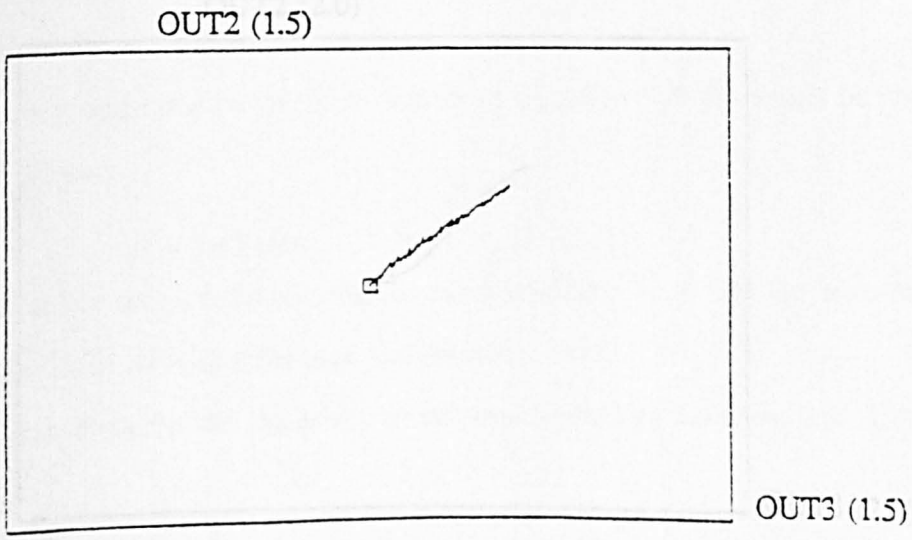


Figure 5.21: Phase plot of Y_{NL} versus X_{LN}

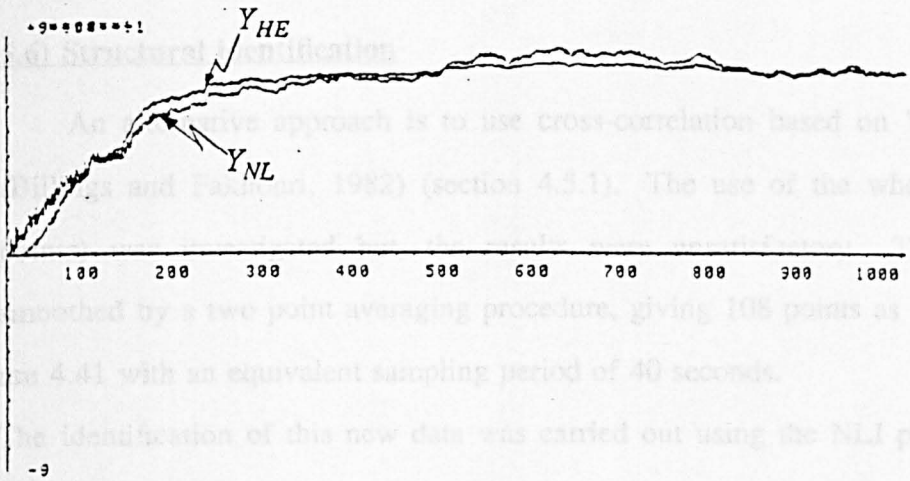


Figure 5.22: Data obtained after optimisation

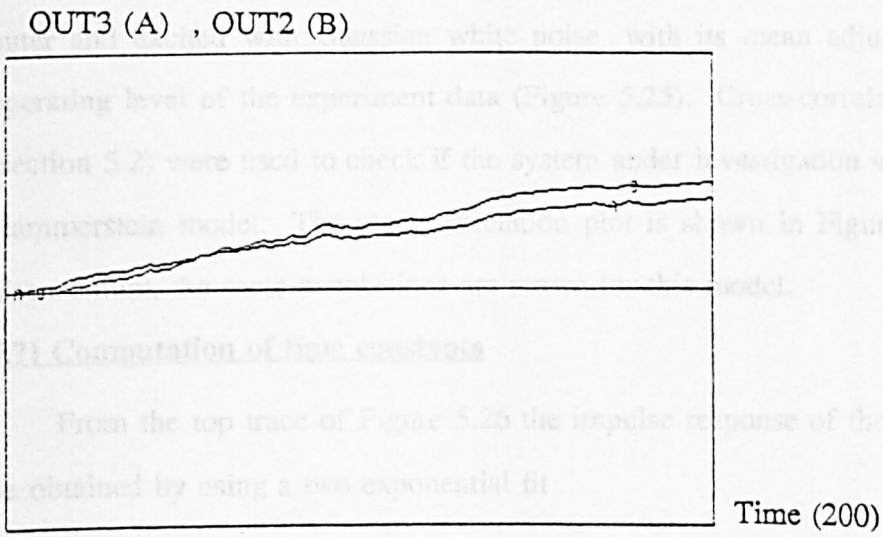


Figure 5.23: Representation of Y_{HE} and X_{LN}

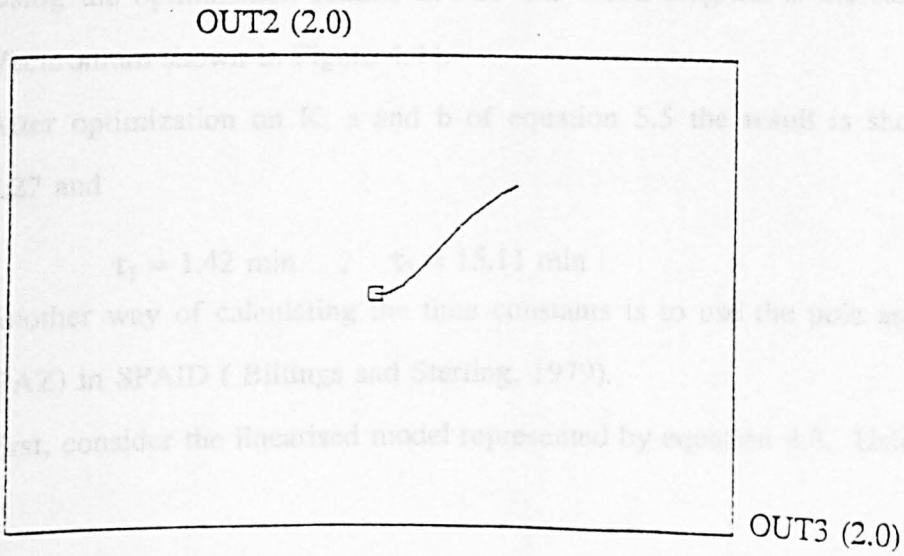


Figure 5.24: Phase plot of Y_{HE} versus X_{LN}

5.6) Structural identification

An alternative approach is to use cross-correlation based on Volterra series (Billings and Fakhouri, 1982) (section 4.5.1). The use of the whole data (218 points) was investigated but, the results were unsatisfactory. The data were smoothed by a two point averaging procedure, giving 108 points as shown in Figure 4.41 with an equivalent sampling period of 40 seconds.

The identification of this new data was carried out using the NLI package . The structural identification was reconsidered by using the model obtained from the smoothed data. The method requires many data points (typically 5000 points). The NARMAX model represented by equation 4.9 was then simulated on a computer and excited with Gaussian white noise with its mean adjusted to suit the operating level of the experiment data (Figure 5.25). Cross-correlation techniques (section 5.2) were used to check if the system under investigation was a Wiener or Hammerstein model. The cross-correlation plot is shown in Figure 5.26. As for Vecuronium, the same conclusions are drawn for this model.

5.7) Computation of time constants

From the top trace of Figure 5.26 the impulse response of the linear part can be obtained by using a two exponential fit

$$\frac{Y(s)}{u(s)} = \frac{K}{(s+a)(s+b)} \tag{5.27}$$

Using the optimization routine in PSI the block diagram is the same as that for Vecuronium shown in Figure 4.44.

After optimization on K, a and b of equation 5.5 the result is shown in Figure 5.27 and

$$\tau_1 = 1.42 \text{ min} \quad , \quad \tau_2 = 15.11 \text{ min}$$

Another way of calculating the time constants is to use the pole and zero option (PAZ) in SPAID (Billings and Sterling, 1979).

First, consider the linearised model represented by equation 4.8. Using the (PAZ)

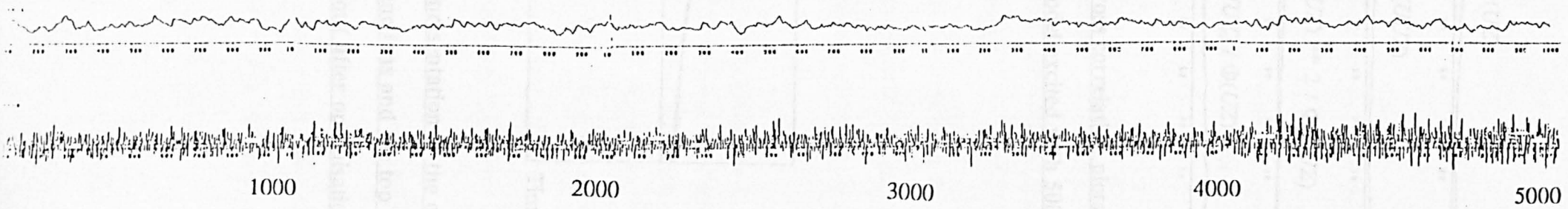


Figure 5.25: Representation of the 5000 point white noise sequence

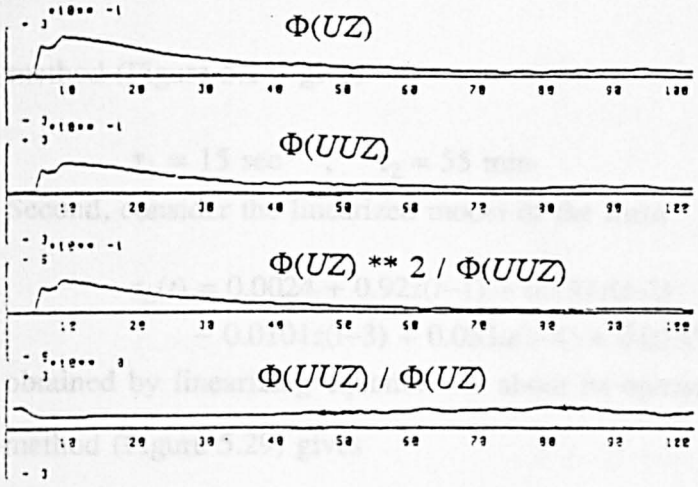


Figure 5.26: Cross correlation plots obtained from the NARMAX model excited with 5000 point white noise sequence

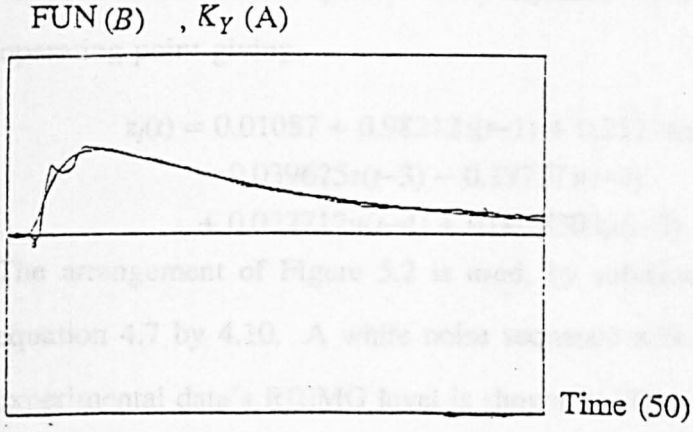


Figure 5.27: Representation of the output of the two exponential transfer functions and the top trace of the cross-correlation plot (after optimisation)

method (Figure 5.28) gives

$$\tau_1 = 15 \text{ sec} \quad , \quad \tau_2 = 55 \text{ min}$$

Second, consider the linearized model of the form

$$z_L(t) = 0.0024 + 0.92z(t-1) + 0.137z(t-2) - 0.0101z(t-3) + 0.033u(t-4) + 0.0135u(t-7) \tag{5.28}$$

obtained by linearizing equation 4.9 about its operating points. The use of (PAZ) method (Figure 5.29) gives

$$\tau_1 = 34 \text{ sec} \quad , \quad \tau_2 = 14.46 \text{ min}$$

The above results are tabulated in table 5.1

5.8) Fitting a physiological structure to the model obtained using the NOI package (Atracurium)

A NARMAX model obtained using the NOI package on the real data can be explored via off-line simulation to fit a physiological structure to this model. The NARMAX model is represented by equation 4.10 and was linearized about an operating point giving

$$z_l(t) = 0.01087 + 0.98212z(t-1) + 0.21716z(t-2) - 0.039625z(t-3) - 0.19737z(t-4) + 0.022712u(t-4) + 0.0028305u(t-5) \tag{5.29}$$

The arrangement of Figure 5.2 is used, by substituting equation 4.8 by 5.27 and equation 4.7 by 4.10. A white noise sequence with its mean adjusted to give the experimental data's RIEMG level is shown in Figure 5.30

Before optimization the plots of Y_{NL} versus X_{LN} are shown in Figures 5.31 and 5.32.

After optimization, the parameters of the Hill equation turned out to be

$$\alpha = 3.20 \quad , \quad D = 0.677$$

Substituting these values into the Hill equation, the Plot of Y_{HE} and Y_{NL} is shown in Figure 5.33 and the plot of Y_{HE} versus X_{LN} is shown in Figures 5.34 and 5.35 respectively.

Poles	
Real	*Imag*
0.000E+00	0.000E+00
0.282E+00	0.000E+00
-0.418E+00	0.000E+00
0.994E+00	0.000E+00

Zeros	
Real	*Imag*
0.572E+00	0.000E+00
0.172E-01	-0.135E-01
0.172E-01	0.135E-01

System Gain=4.811861

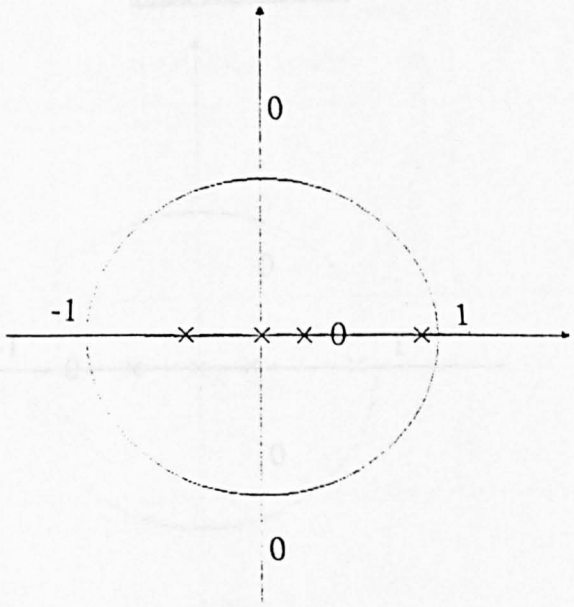


Figure 5.28: Poles and zeroes representation

Poles		Zeros	
Real	*Imag*	*Real*	*Imag*
0.000E+00	0.000E+00	0.370E+00	-0.541E+00
-0.342E+00	0.000E+00	0.370E+00	0.541E+00
0.310E+00	0.000E+00	-0.740E+00	0.000E+00
0.955E+00	0.000E+00		

System Gain=1.137918

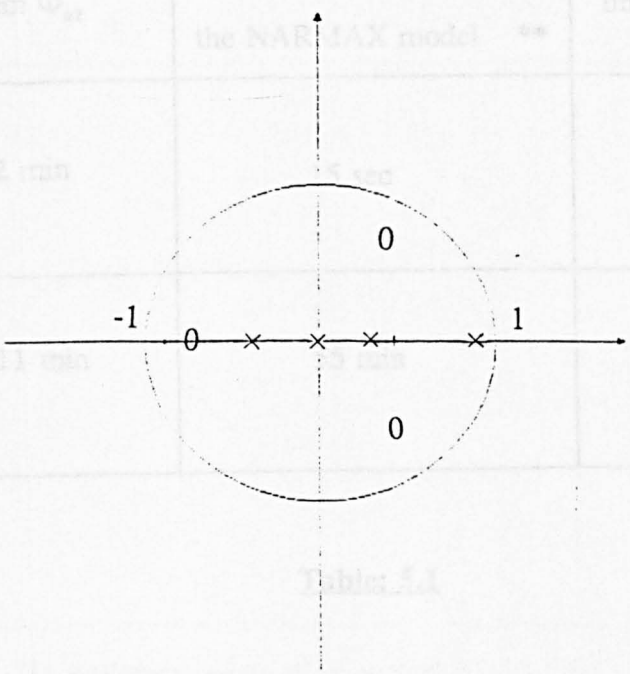


Figure 5.29: Poles and zeroes representation

	Fit of second order system Φ_{uz} *	The whole 218 points being nonlinear and then linearized the NARMAX model **	The smoothed data (108 points) and then linearized the NARMAX model ***
τ_1	1.42 min	15 sec	34 sec
τ_2	15.11 min	55 min	14.46 min

Table: 5.1

* Using NLI package on 108 points

** Using NLI package on 218 points

*** Using NLI package on the smoothed data (108 points)

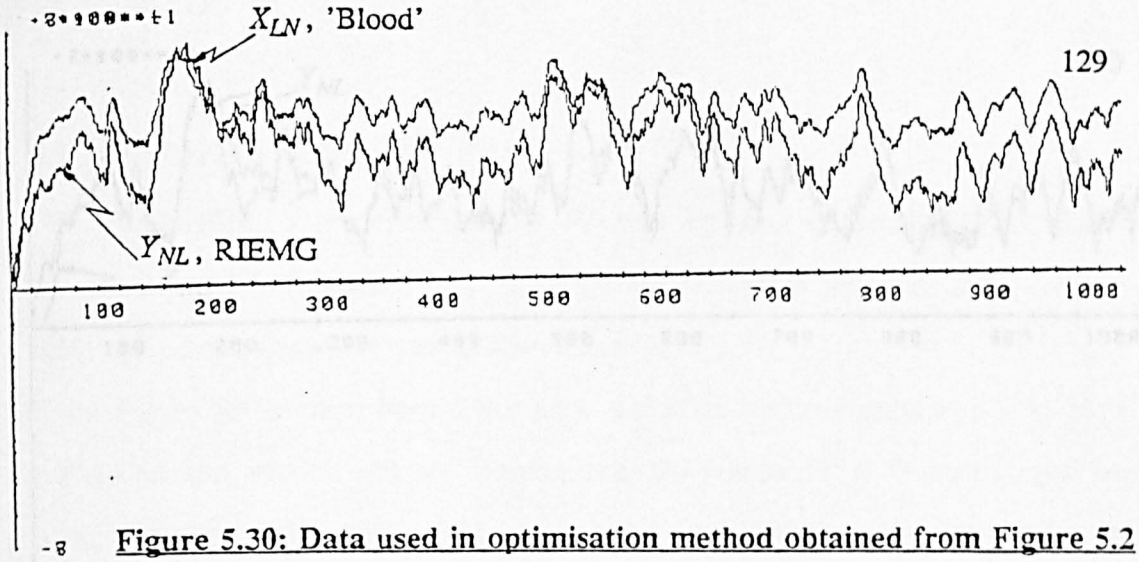


Figure 5.30: Data used in optimisation method obtained from Figure 5.2

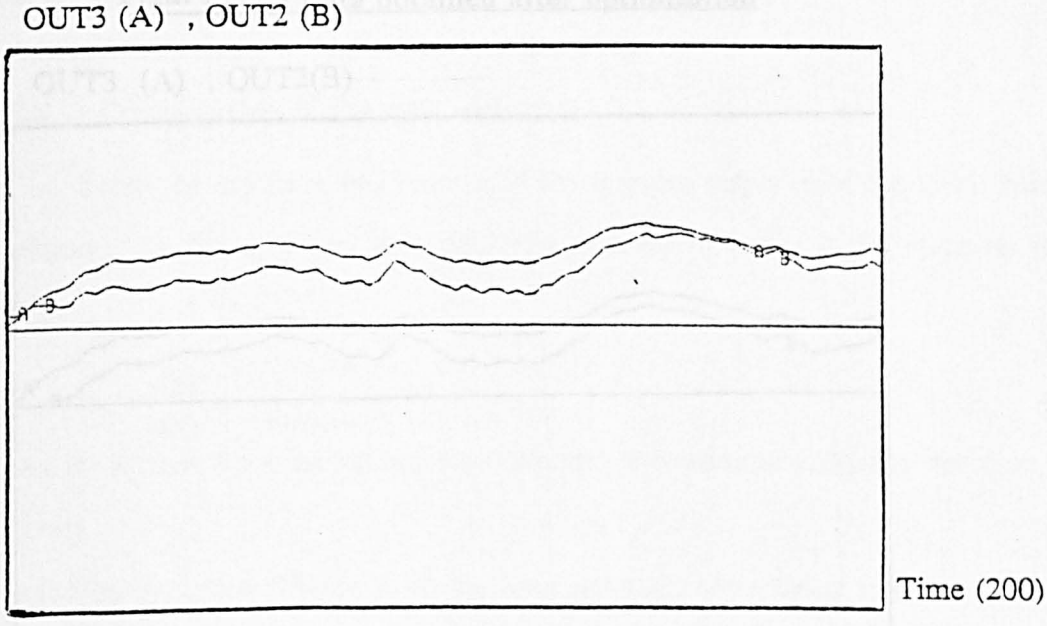


Figure 5.31: Representation of Y_{NL} and X_{LN}

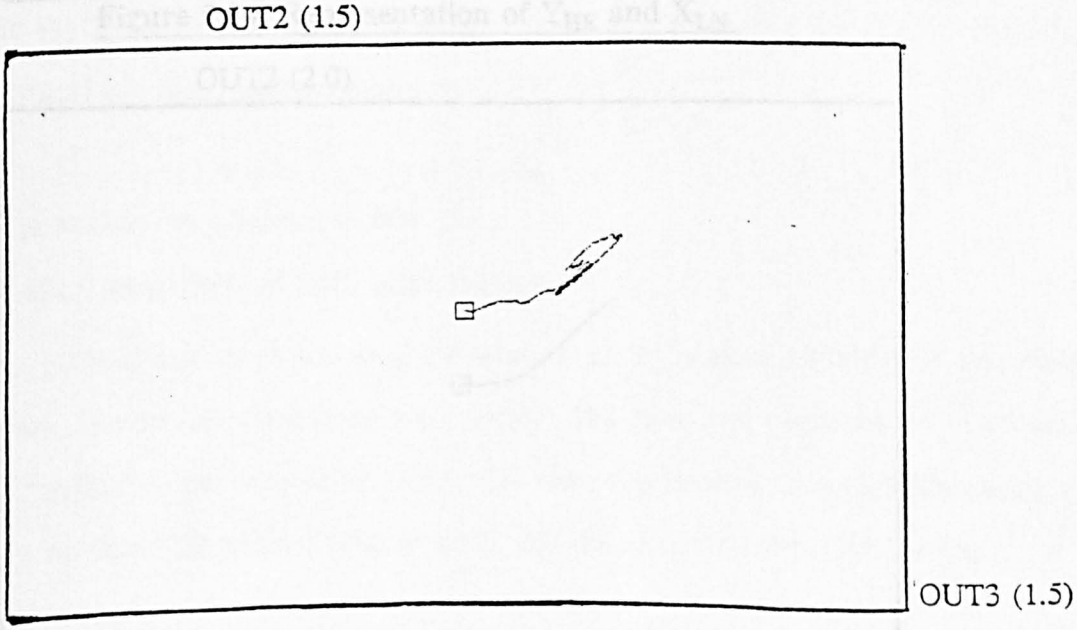


Figure 5.32: Phase plot of Y_{NL} versus X_{LN}

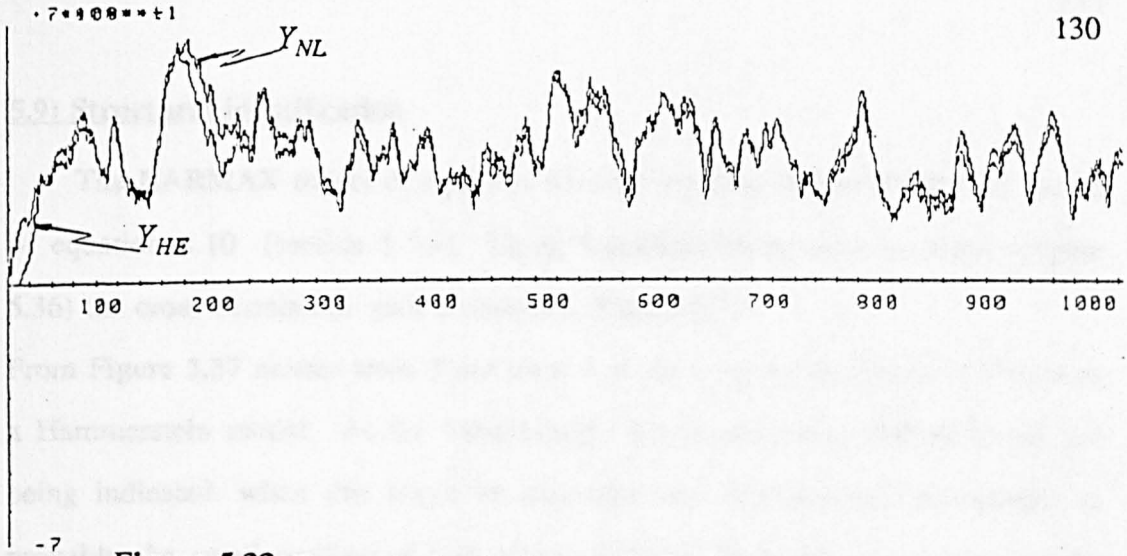
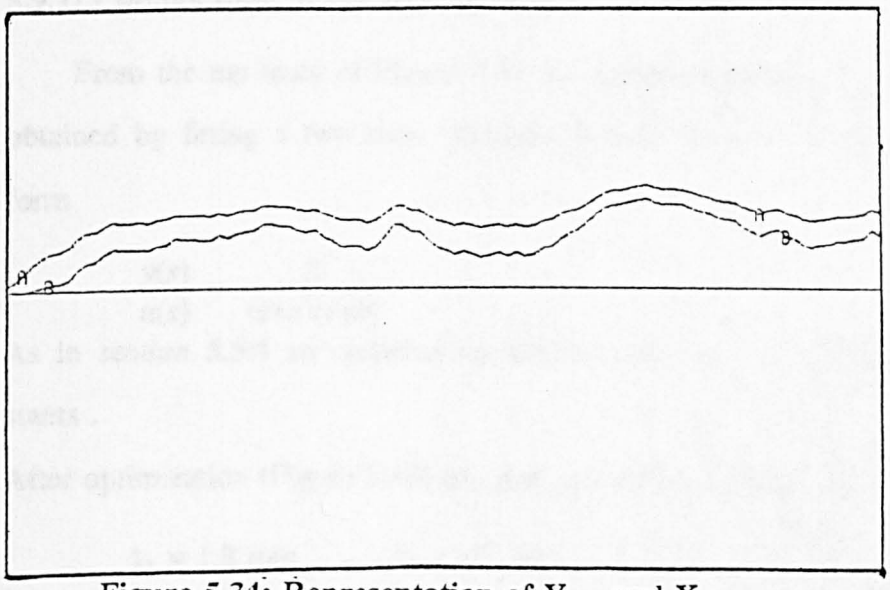


Figure 5.33: Data obtained after optimisation

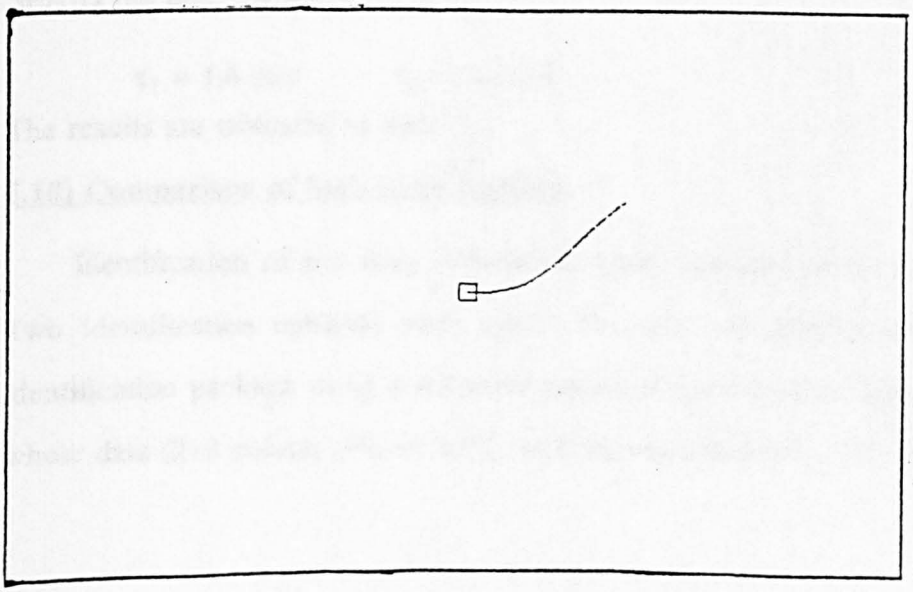
OUT3 (A) , OUT2(B)



Time (200)

Figure 5.34: Representation of Y_{HE} and X_{LN}

OUT2 (2.0)



OUT3 (2.0)

Figure 5.35: Phase plot of Y_{HE} versus X_{LN}

5.9) Structural identification

The NARMAX model of equation 4.9 was replaced by the NARMAX model of equation 4.10 (section 5.5.3). Using Gaussian white noise as input (Figure 5.36) the cross-correlation plot is shown in Figure 5.37.

From Figure 5.37 neither trace 3 nor trace 4 verifies the conditions of a Wiener or a Hammerstein model. As for Vecuronium, the reason for a Wiener model not being indicated, when this might be expected from physiological knowledge, is probably the small number of data points obtained from the experiment and the non-optimality of the shape of the input signal.

5.9.1) Computation of the time constants

From the top trace of Figure 5.37 the impulse response of the linear part was obtained by fitting a two time constants transfer function to the response of the form

$$\frac{y(s)}{u(s)} = \frac{K}{(s+a)(s+b)} \tag{5.30}$$

As in section 5.5.4 an optimization routine was used to calculate the time constants .

After optimization (Figure 5.38) the time constants were found to be

$$\tau_1 = 1.3 \text{ min} \quad , \quad \tau_2 = 15 \text{ min}$$

If the linearized model is considered, equation 5.8, using the (PAZ) method in SPAID the time constants become

$$\tau_1 = 1.6 \text{ min} \quad , \quad \tau_2 = 22 \text{ min}$$

The results are tabulated in table 5.2

5.10) Comparison of both identifications

Identification of the drug Atracurium under manual control was performed. Two identification methods were used. The first one consisted of a nonlinear identification package using a recursive extended least squares algorithm using the whole data (218 points) (Figure 5.11), and the smoothed data (108 points)

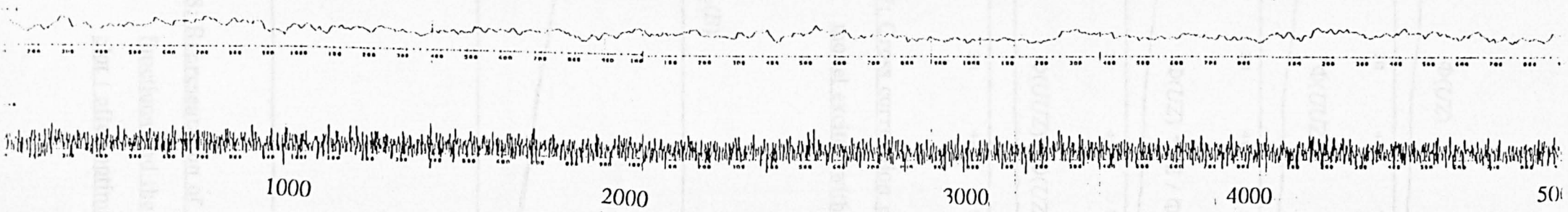


Figure 5.36: Representation of the 5000 point white noise sequence

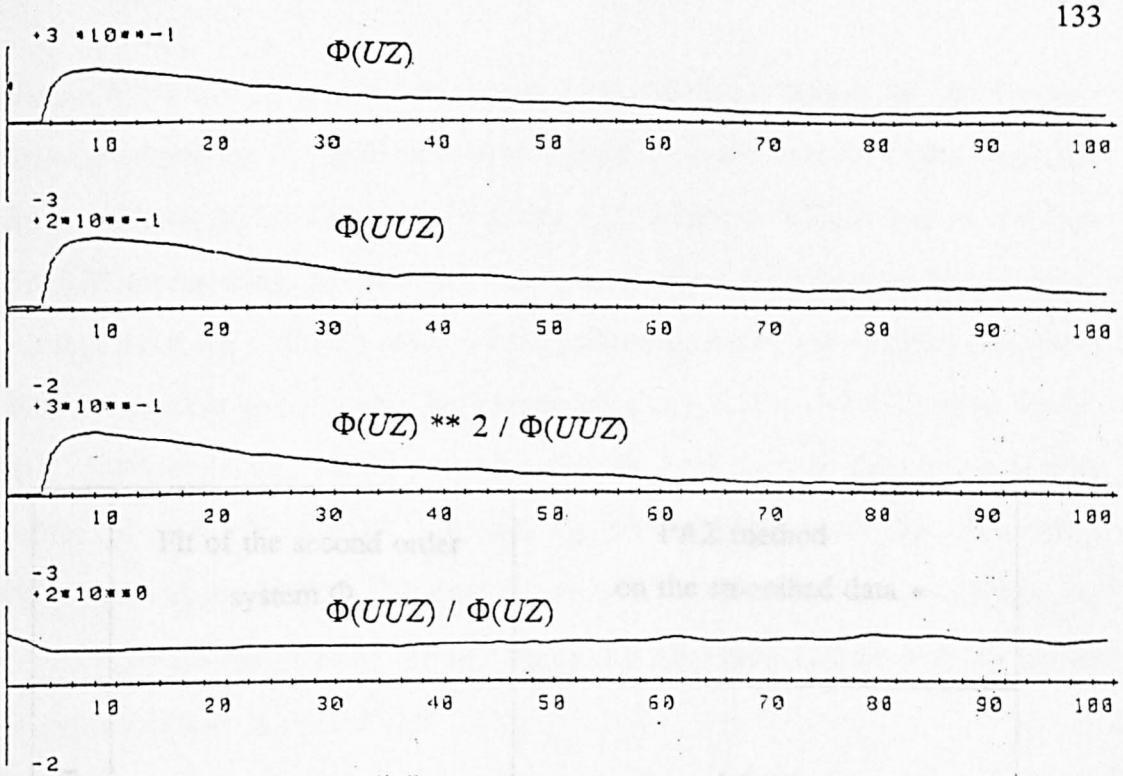


Figure 5.37: Cross correlation plots obtained from the NARMAX model excited with 5000 point white noise sequence

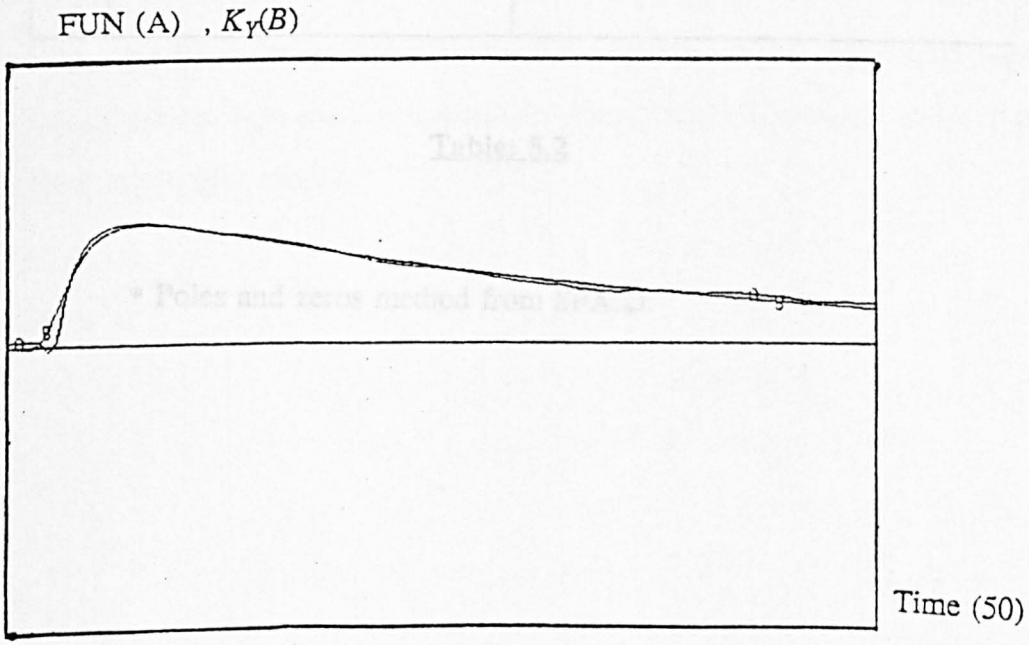


Figure 5.38: Representation of the output of the two exponential transfer functions and the top trace of the cross-correlation plot (after optimisation)

	Fit of the second order system Φ_{uz}	PAZ method on the smoothed data *
τ_1	1.3 min	1.6 min
τ_2	15 min	22 min

Table: 5.2

* Poles and zeros method from SPAID.

(Figure 5.26) and the second one consisted of a nonlinear orthogonal identification package employing the prediction error algorithm on the smoothed data only (108 points) (Figure 5.35). Comparing Figures 5.11, 5.26 and 5.35, it is clear that Figure 5.35 representing the NOI technique demonstrates the effectiveness of the algorithm since the predicted output of the estimated model was virtually coincident with the process output. Also the parameters ($\alpha = 3.20$, D=0.67) found for the Hill equation via the model obtained using the NOI package compare well with values of $\alpha = 2.6$, D=0.4 obtained by traditional methods for Atracurium (Weatherly et al, 1983). The time constants $\tau_1 = 1.3$ min and $\tau_2 = 15$ min obtained from the model using the NOI technique are commensurate with the known pharmacokinetics of Atracurium.

The reason for not giving a good structure by using NLI method is due to the algorithmic limitation of the package. Hence, using a collection of nonlinear identification methods currently available it has been possible to obtain parameters for a model which describes a dose/response of muscle relaxant drug administration. Such a NARMAX model could be used for the design of improved controller algorithms for on-line drug infusion regimes. These algorithms vary from simple three-term to minimum variance optimal control strategies and will be discussed in the next chapter.

CHAPTER 6

Simulation-Based Control Of Muscle Relaxant Administration

6.1) Introduction

In the past few years, automatic control has experienced a large surge of interest in all fields of engineering. Systems requiring to be controlled automatically are becoming larger, more powerful, widespread and diverse, emcompassing ventilation control, combustion control in high temperature furnaces, pH control in chemical plants and automated operation of power plants, aircraft, ships and radio circuits. New applications for automatic control are continually coming to the fore.

In recent years, automatic control has found new applications and has become a well established theory in the field of biomedical engineering. Linkens (1979) shows many applications of automatic control covering several areas of biomedical systems (lung parameter estimation, digestive tract, eye-and-ear control, heart beat arythmia).

In the field of anaesthesia (Brown et al, 1980, Asbury et al, 1980, Linkens et al, 1981, Cass et al, 1976, Shepphard et al, 1979) demonstrate the superiority of feedback control over the conventional manual methods.

It has been mentioned in chapter 2 that the manual control of muscle relaxant administration tends to proceed by over-paralysis. To overcome this problem, antagonist drugs are administered to reverse the neuromuscular block. Sometimes, too large a dosage may not be desirable for the patient as this gives rise to post-operative complications. The magnitude of the dosage and the maintaining of correct levels of relaxation are the anaesthetist's main concerns in the control of the muscle relaxant administration. Medical staff would be relieved from the task of constant observation of the bedside monitors and frequent adjustments of the

infusion rate to maintain control by the introduction of feedback.

The work reported in this chapter concerns the on-line control of muscle relaxant dynamics of both the drugs: Vecuronium and Atracurium. It is evident, from the control literature, that greater emphasis has been placed on linear systems. The area of adaptive control is no exception. This imbalance, can of course, be attributed to the diversity and complexity of nonlinear systems.

However, many important control problems are known to be inherently nonlinear. Therefore, it is reasonable to believe that better control would result by accurately identifying a nonlinear model for the system and incorporating this knowledge into the structure of the control law. Hence nonlinear identification was carried out and NARMAX models were obtained for both the above drugs (chapter 4).

Three different strategies will be considered in the present chapter to control the NARMAX models for Vecuronium and Atracurium. Section 6.2 will consider the most widely used controller in control system design. It is the three-term or PID (Proportional, Integral and Derivative) controller algorithm, and is given by

$$u(t) = K\left[e(t) + \frac{1}{T_i}\int_0^t e(t)dt + T_d\frac{de(t)}{dt}\right] \tag{6.1}$$

where K is the proportional gain, T_i is the integration time, T_d is the derivative time, u(t) and e(t) represent the control and error signals respectively. Section 6.3, deals with the application of the general minimum variance self-tuning controller developed by Clarke and Gawthrop (1975, 1979) and Clarke et al, (1975), which was originally introduced by Astrom and Wittemark (1973). This controller is based on the minimization of a cost function that allows for set-point tracking of continuously varying demand signals and includes control effort weighting to avoid generation of excessive control signals. In section 6.4 a generalised predictive controller (GPC) algorithm due to Clarke et al, (1987) is used to control the NARMAX models of both muscle relaxants. Linkens and Mahfouf (1988) used

this algorithm successfully to control an overall nonlinear model of the drug Atracurium in which the pharmacokinetics of the drug were modelled by a third order transfer function while the pharmacodynamics (nonlinear part) were represented by a Hill equation.

6.2) Design of the PID controller

The PID controller has a forty-year history as the work horse of the process industries. Most industrial loops are controlled by discrete versions of the basic PID controller. With the surge of recent years in the area of adaptive controllers, tuning has been a major preoccupation amongst control theorists. The tuning scheme was divided into two categories: off-line and on-line.

In off-line methods, controller parameters are calculated applying the early scheme of Ziegler and Nichols (1942). Mcgregor et al (1975) used a method based on identification, followed by numerical optimization of the control parameters.

In contrast with off-line methods, on-line methods attempt to change controller settings as data are received. Work in adaptive control falls into the on-line category (self-tuning regulators, model reference adaptive control). In this section, an off-line method is used to design PID controllers for the NARMAX models of the muscle relaxant dynamics.

Equation 6.1 can be expressed in discrete time as

$$\frac{u(t)}{e(t)} = K_p + \frac{K_i}{(1 - z^{-1})} + K_d(1 - z^{-1}) \tag{6.2}$$

or

$$u(t) = u(t-1) + P_0 e(t) + P_1 e(t-1) + P_2 e(t-2) \tag{6.3}$$

where

$$\left.\begin{aligned} P_0 &= K_p + K_i + K_d \\ P_1 &= -K_p - 2\,K_d \\ P_2 &= K_d \end{aligned}\right\} \tag{6.4}$$

Using the simplex method (Nelder and Mead, 1965), the off-line tuning of the

parameters P_0 , P_1 , P_2 (equation 6.3) was formulated as an optimization problem in which the cost function to be minimised was the integral of the squared error. The closed loop configuration, inclusive of the PID controller whose parameters are to be optimised is shown in Figure 6.1 below.

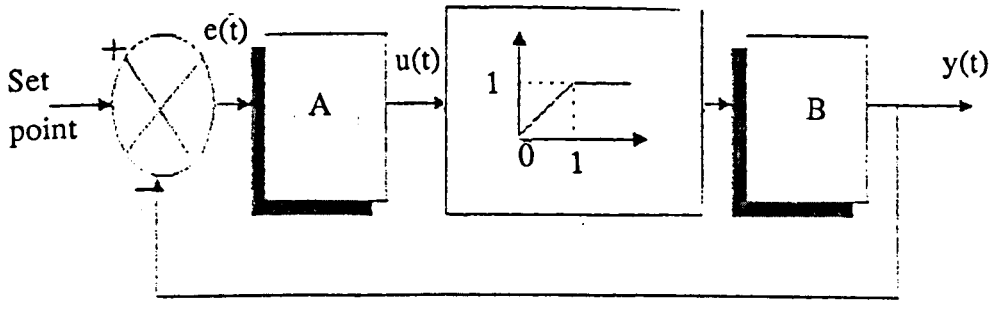


Figure 6.1: Three-term control of the NARMAX models

With reference to Figure 6.1 , block A represents the PID controller and B the relaxant model: either that of Vecuronium or Atracurium . Note that in either case the control signal is limited between 0 and 1 prior to its application to the system.

1) Case of Vecuronium model

With the above set-up, the optimisation procedure yielded the following PID parameters

$$P_0 = 0.95 \quad , \quad P_1 = 0.1 \quad , \quad P_2 = 2.8$$

With the controller parameters fixed to these values, the responses of the system to square wave demands of levels (0.7 and 0.4) and (0.7 and 0.8) are shown in Figures 6.2 and 6.3 respectively.

2) Case of Atracurium model

Two situations are considered for this case which are:

a) that of the drug model obtained using the NLI package (equation 5.4) and

b) that of the model obtained via the NOI package (equation 5.7).

The controller parameters for these cases turned out to be

P_0 =1.5 P_1 =0.005 , P_2 =3.0

P_0 =1.95 P_1 =0.15 , P_2 =2.3

respectively

The system responses for both the above situations to the same demand signals as those for the Vecuronium case are shown in

a) Figures 6.4 and 6.5 and

b) Figures 6.6 and 6.7

respectively

The responses of Figures 6.2, 6.3, 6.4, 6.5, 6.6 and 6.7 demonstrate fairly satisfactory PID control of the Vecuronium and Atracurium NARMAX models. However, these responses were obtained with the model parameters fixed to those values for which controller parameter optimisations were performed in the absence of any noisy disturbances. So, if the systems parameters and/or time delay were to vary, which is the case in practice (Menad, (1984) discusses this aspect and gives the ranges within which these quantities do vary in practical situations), it is to be expected that the performance of these controllers will deteriorate gradually the further these quantities get from their 'nominal' values i.e those values upon which controller optimisation is based.

Further performance degradation is expected upon the introduction of noisy disturbances even if the derivative term in the PID controller were absent.

Despite the above problems associated with fixed term feedback control of relaxant administration, Linkens et al, (1982), Brown et al, (1980), Menad, (1984), Rametti, (1985) among other researchers have underlined its superiority to that of the conventional manual control method.

An even better control strategy would be one in which an acceptable performance

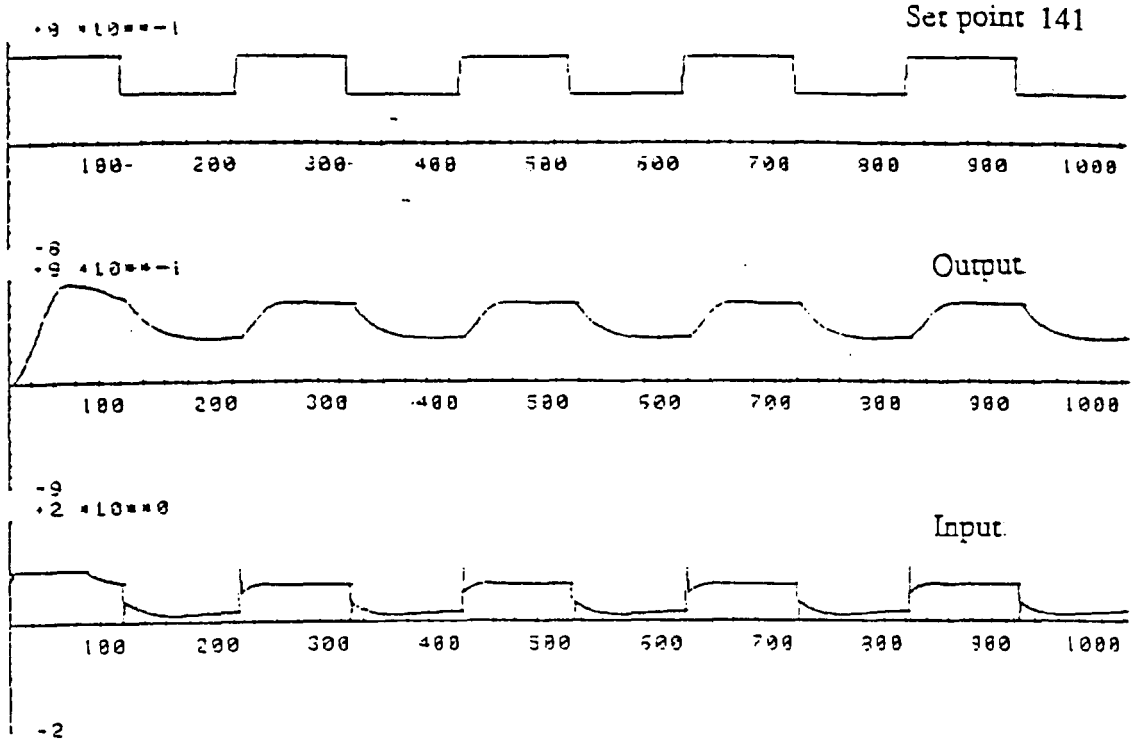


Figure 6.2: Control of the Vecuronium model (NLI) using optimized PID

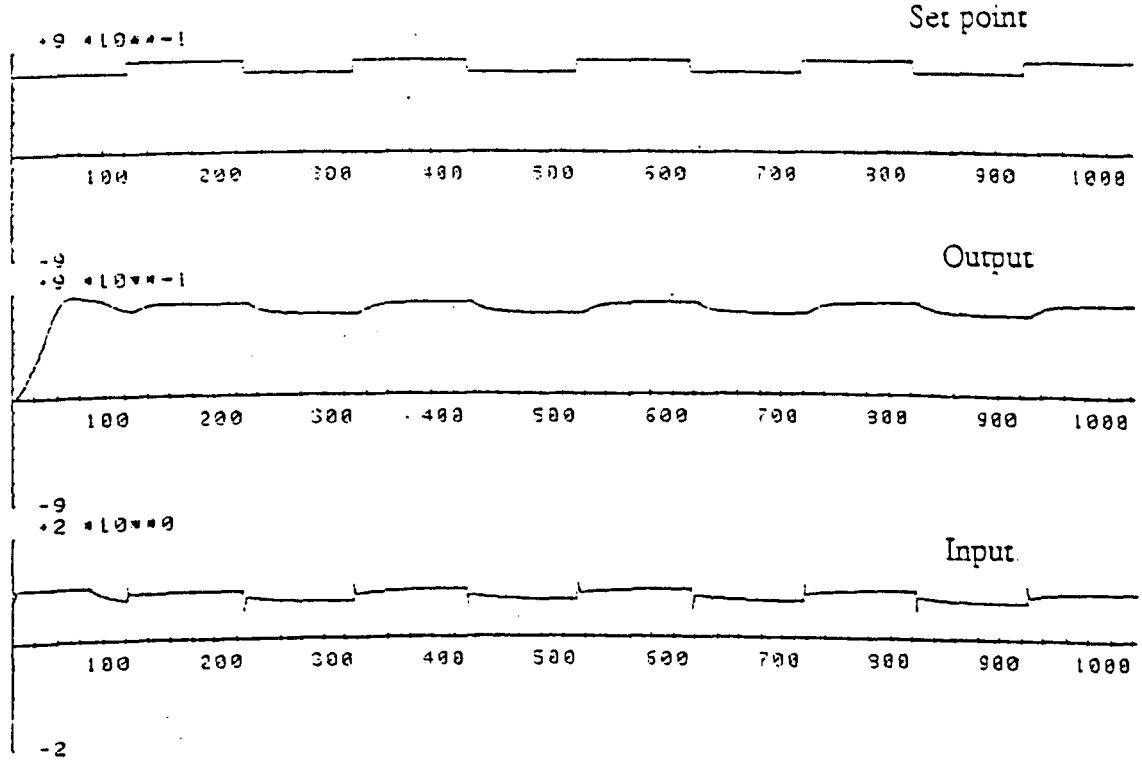


Figure 6.3: Control of the Vecuronium model (NLI) using optimized PID

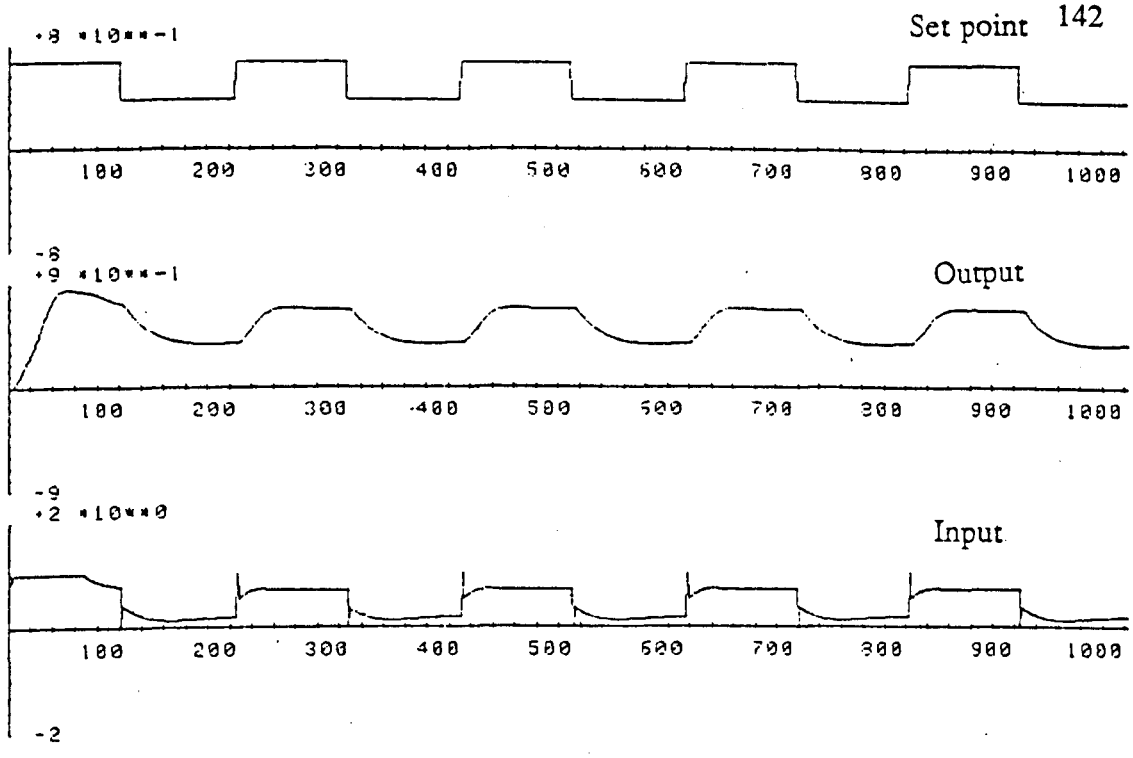


Figure 6.2: Control of the Vecuronium model (NLI) using optimized PID

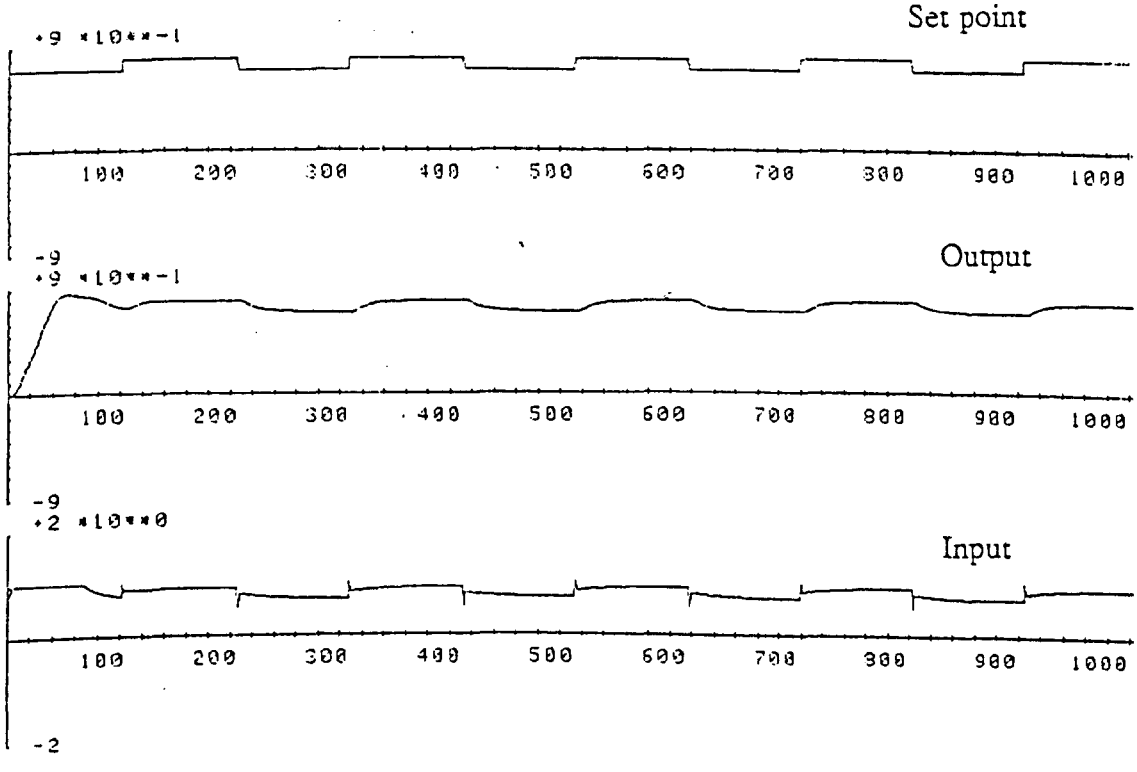


Figure 6.3: Control of the Vecuronium model (NLI) using optimized PID

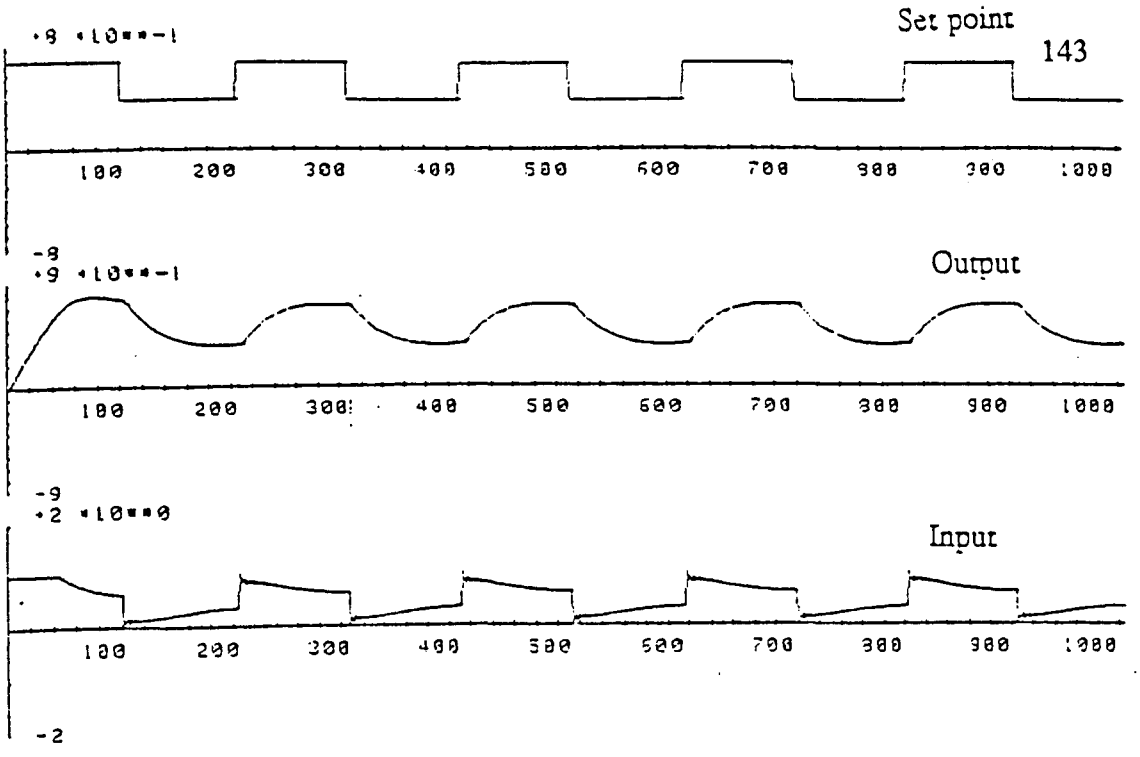


Figure 6.4: Control of the Atracurium model (NLI) using optimized PID

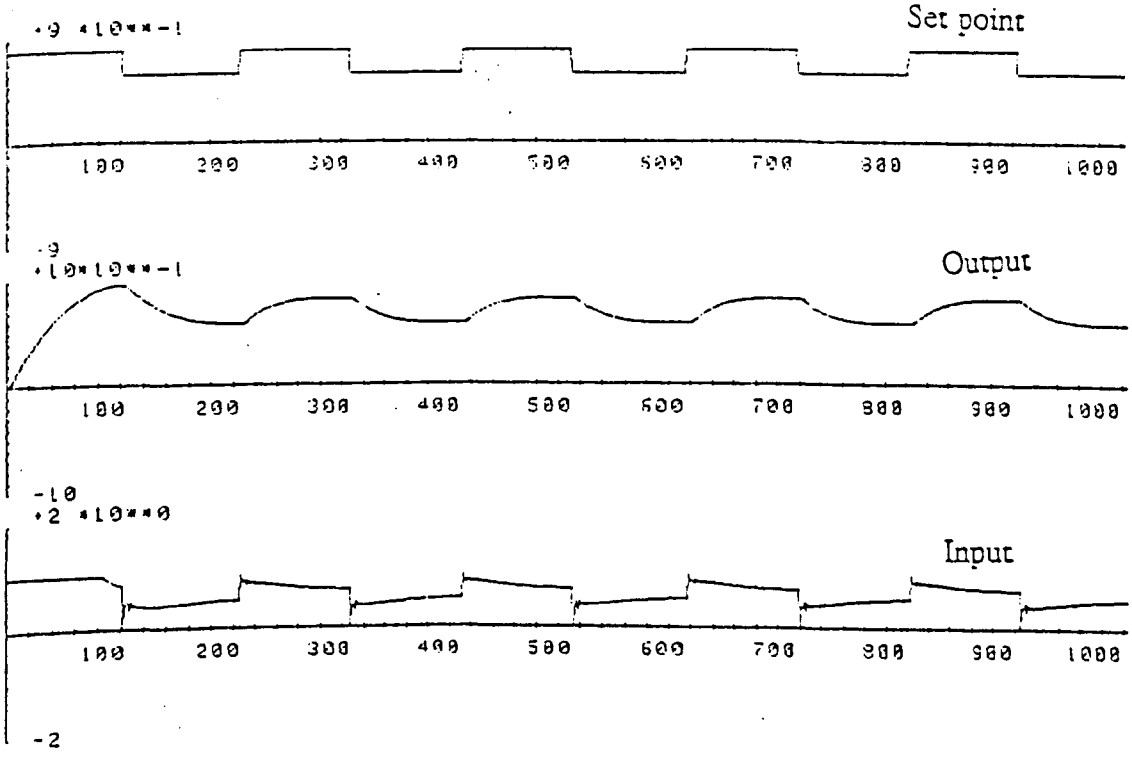


Figure 6.5: Control of the Atracurium model (NLI) using optimized PID

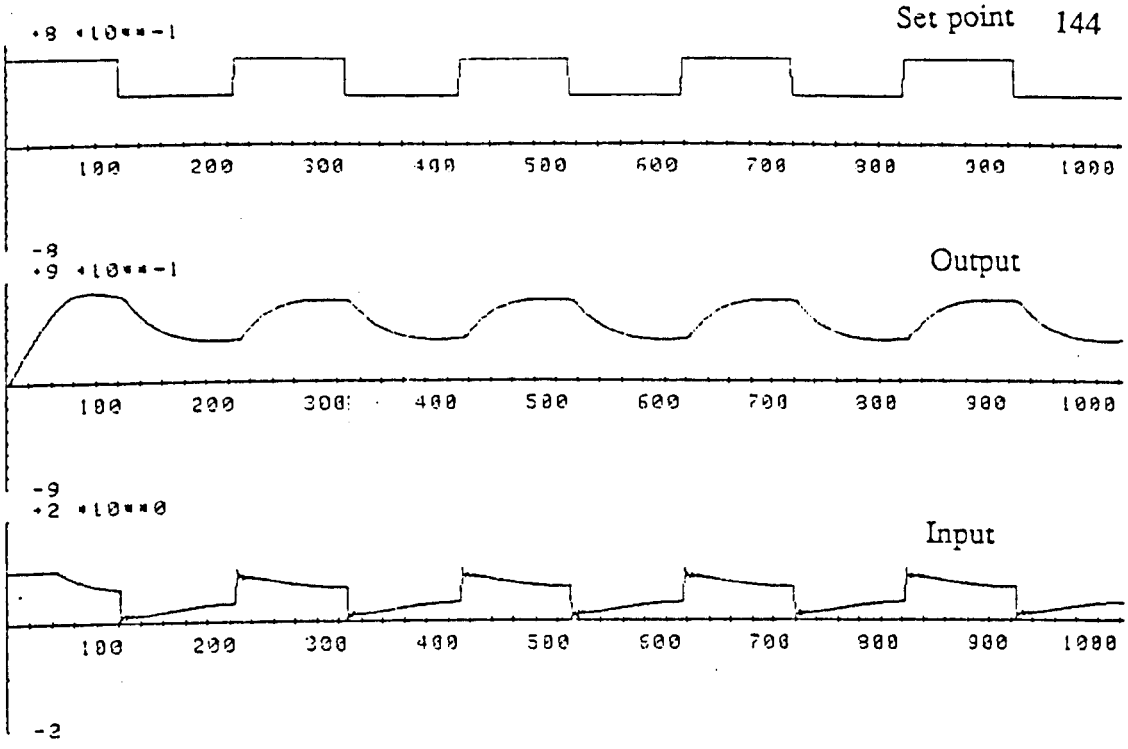


Figure 6.4: Control of the Atracurium model (NLI) using optimized PID

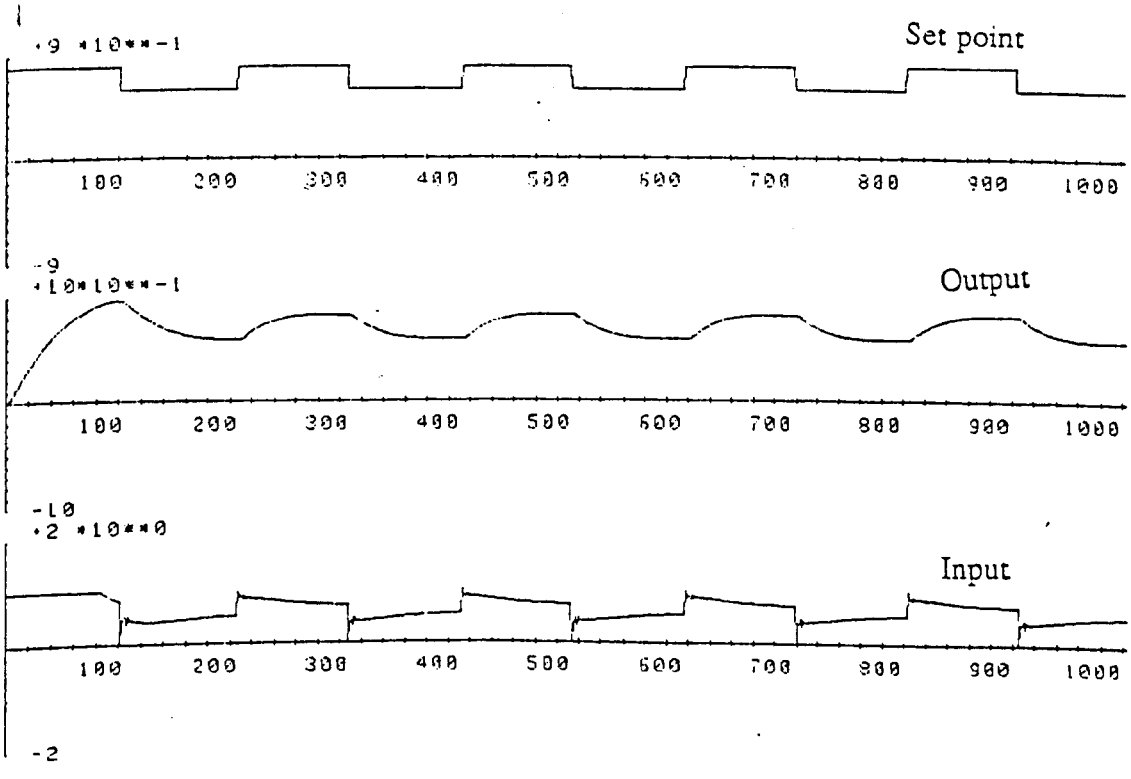


Figure 6.5: Control of the Atracurium model (NLI) using optimized PID

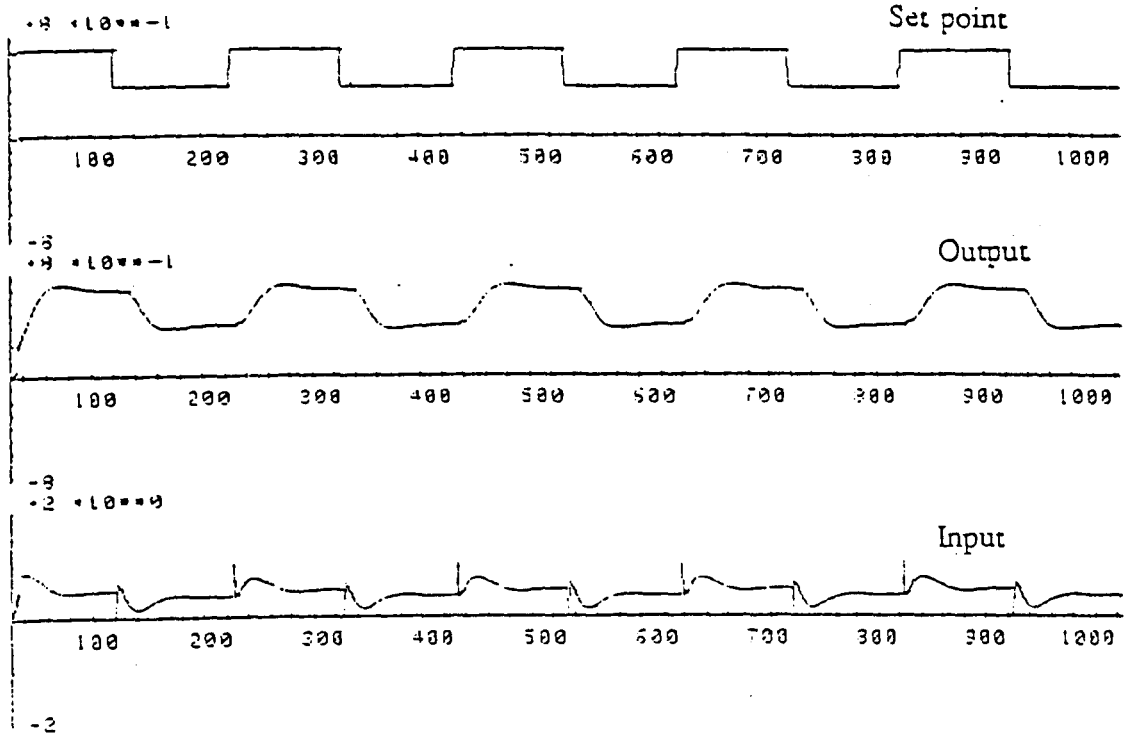


Figure 6.6: Control of the Atracurium model (NOI) using optimized PID

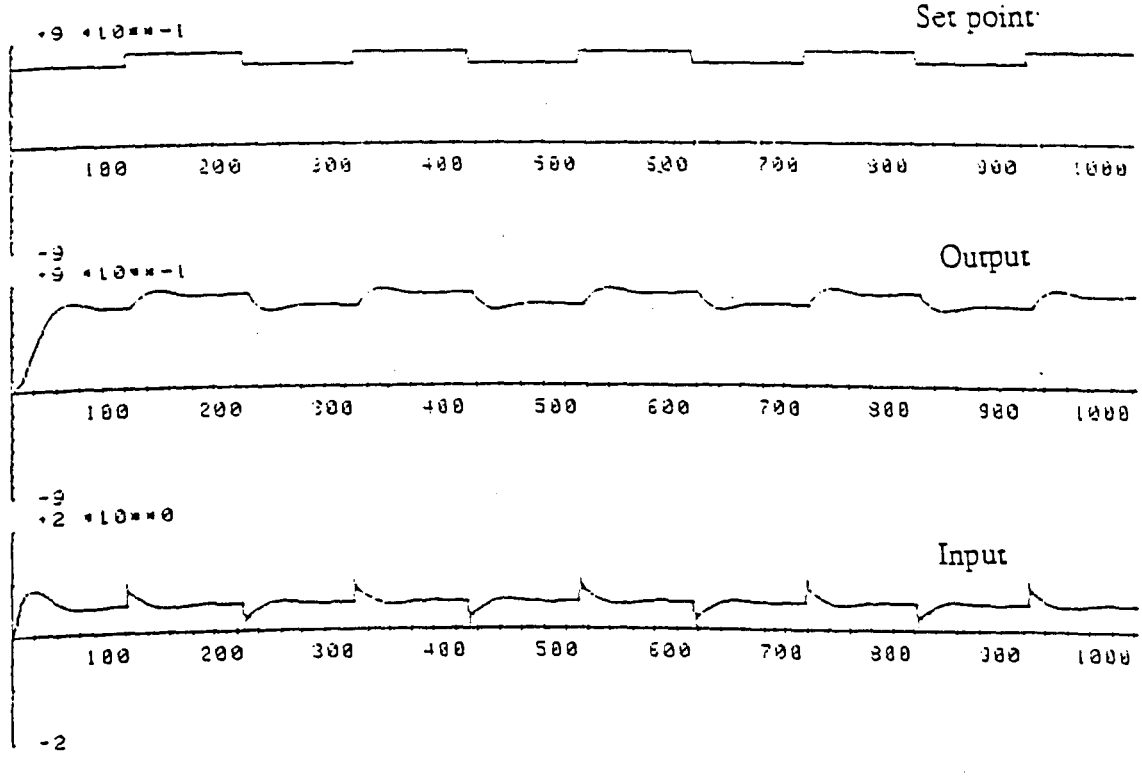


Figure 6.7: Control of the Atracurium model (NOI) using optimized PID

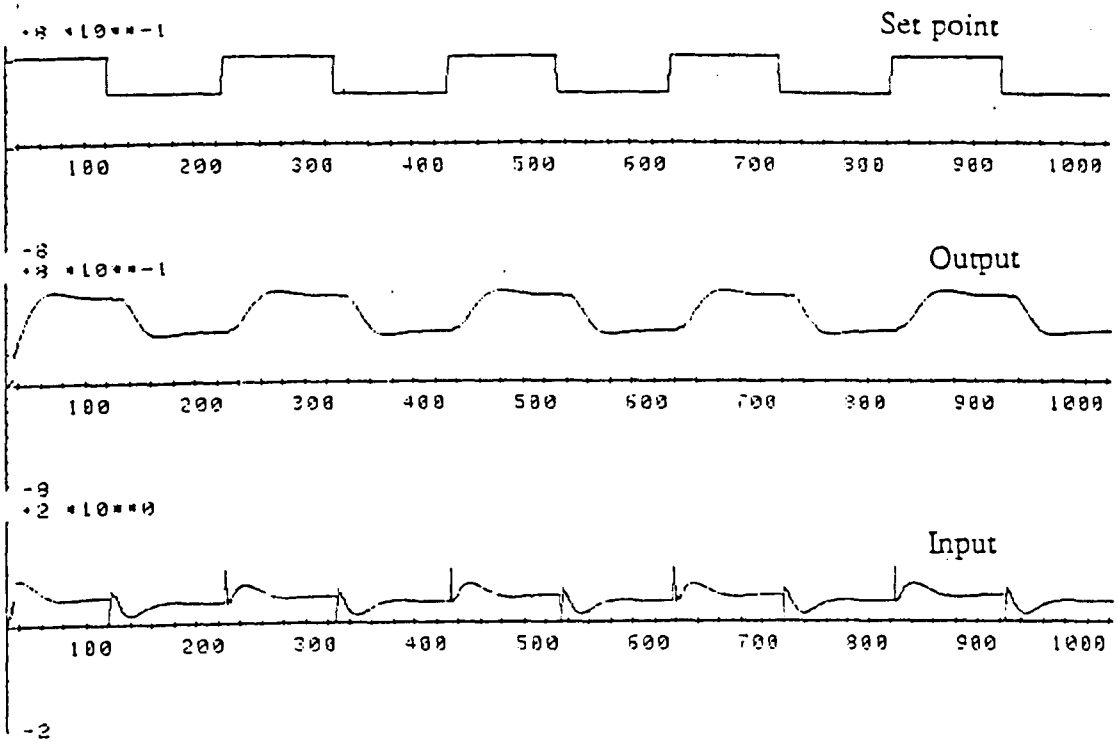


Figure 6.6: Control of the Atracurium model (NOI) using optimized PID

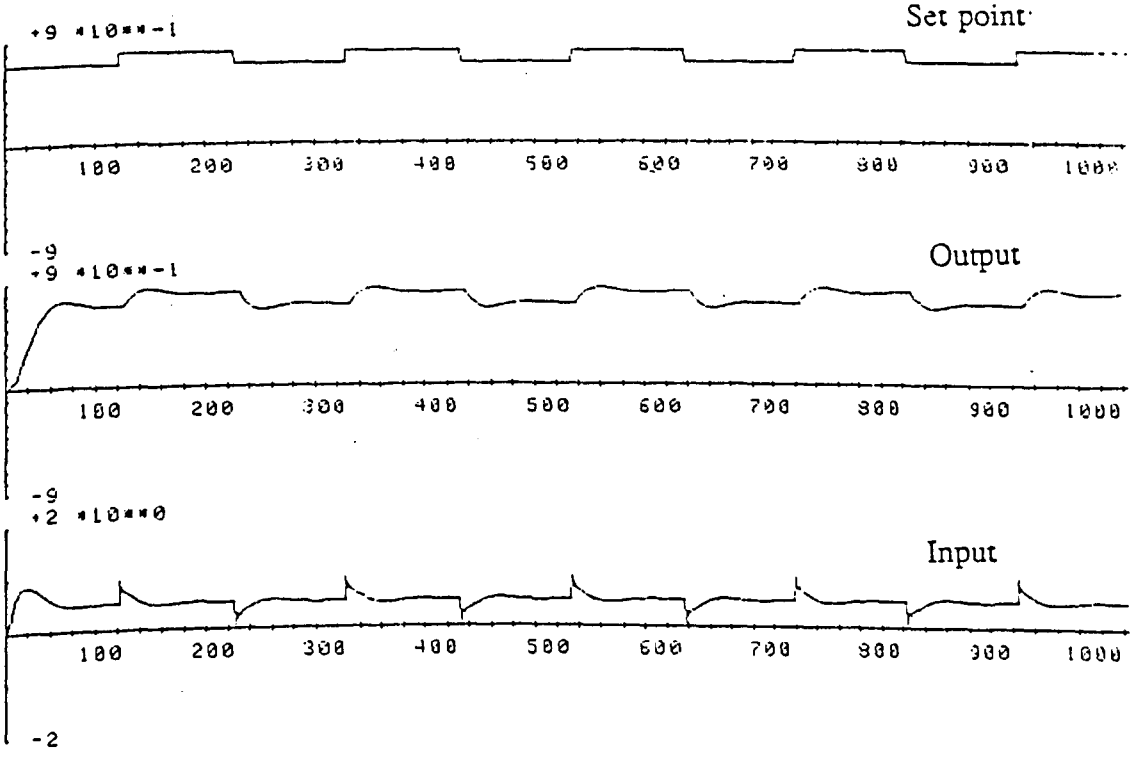


Figure 6.7: Control of the Atracurium model (NOI) using optimized PID

is maintained even under varying parameter and noisy (providing the signal to noise ratio is high enough) conditions.

6.3) General minimum variance (GMV)

Superior control (to that of the fixed-term controllers) can be achieved with a class of controllers known as self-tuning controllers (Clarke and Gawthrop, 1975, 1979), Wellstade et al, 1979, Astrom and Wittermark, 1973).

Such controllers combine the identification and control phases on-line so that any system parameter changes are tracked and the controller parameters changed accordingly to achieve whatever objective constitutes the basis of the design (e.g minimisation of some performance index (Clarke and Gawthrop, 1975, 1979) or assignement of closed loop poles (Wellstead et al, 1979)).

Compared with the standard three-term controller, adaptive controllers are relatively sophisticated and should not be used if a simple constant parameter controller will do the job.

The general structure of the self-tuning controller is shown in Figure 6.8.

A self-tuning regulator consists of three parts.

(1) A feedback law that would give good control if the parameters of the process were known.

(2) A recursive estimator which monitors the inputs and the outputs and estimates the unknown parameters.

(3) The controller design algorithm which uses the estimates in (2) to update the control law in (1).

Menad (1984) and Linkens et al (1985) reported the successful application of pole placement self-tuning control (Wellstead et al, 1979) and Denai (1988) that of self-tuning PID control (Kim and Choi, 1987, Kraus et Myrom, 1984, Gawthrop, 1986) in muscle relaxant administration. The following sections of this chapter introduce the general minimum variance (Clarke and Gawthrop, 1975, 1979) and

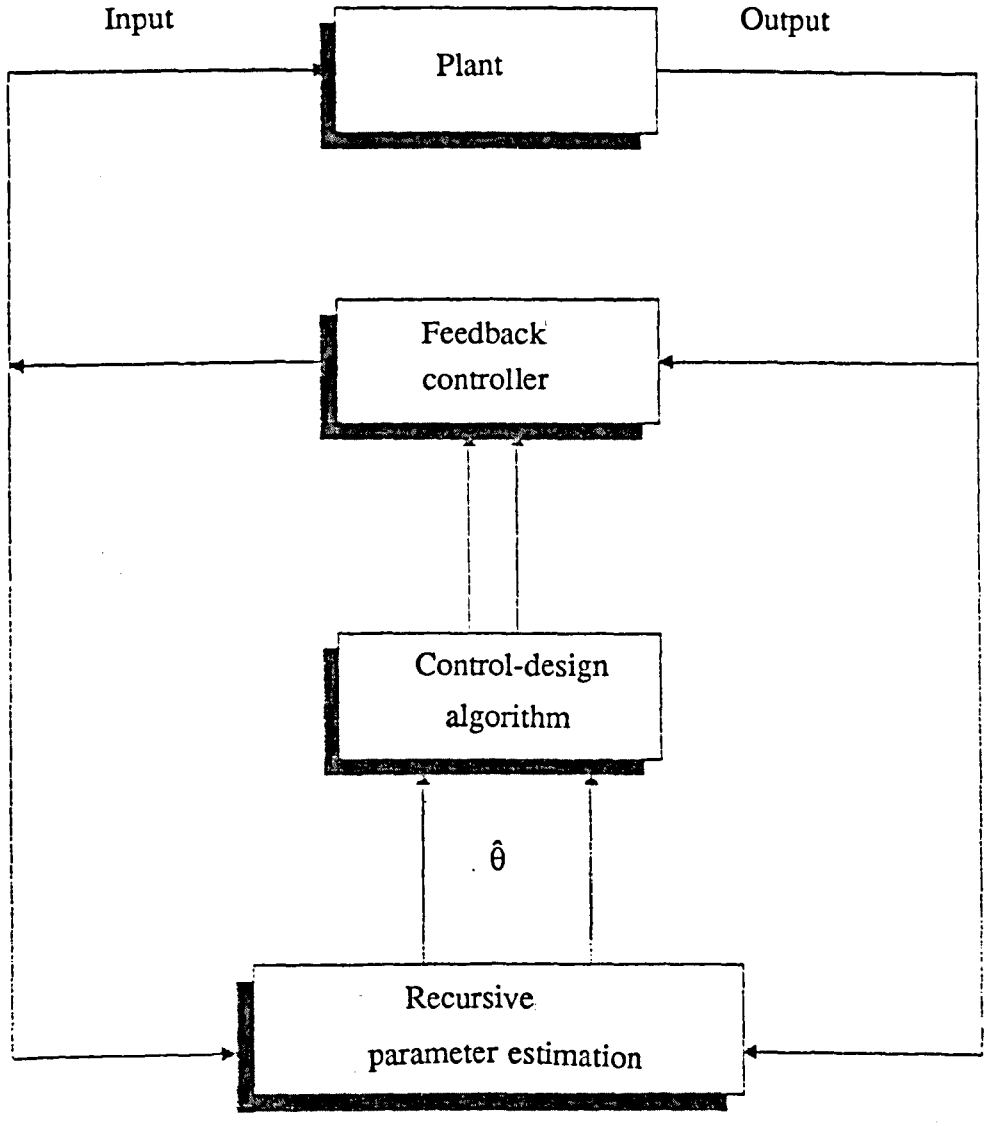


Figure 6.8: Self-tuning sequence

consider its application to the NARMAX models of the drugs Vecuronium and Atracurium.

6.3.1) Minimum variance controller

Consider the system illustrated in Figure 6.9. The system model can be expressed as

$$y(t) = \frac{B(z^{-1})}{A(z^{-1})}u(t-k) + \frac{C(z^{-1})}{A(z^{-1})}\zeta(t) \tag{6.5}$$

or

$$y(t+k) = \frac{B}{A}u(t) + \frac{C}{A}\zeta(t+k) \tag{6.6}$$

or still

$$Ay(t+k) = Bu(t) + C\zeta(t+k) \tag{6.7}$$

where the argument z^{-1} of the polynomials $A(z^{-1}), B(z^{-1})$ and $C(z^{-1})$ has been dropped, and where $k \geq 1$ such that $b_0 = 0$ and the roots of $C(z^{-1})$ are assumed to be within the unit circle.

Consider the control law which minimize the variance of the output

$$I = E[Y^2(t+k)] \tag{6.8}$$

The choice of this criterion originates from a quality control argument where it is desirable that the fluctuations or variance in the plant output around the operating point should be as small as possible.

Since it is the noise which causes the variations in the system output, consider initially the noise term in equation 6.6

$$\frac{C}{A}\zeta(t+k) = \frac{1 + c_1z^{-1} + \ldots\ldots + C_nz^{-n}}{1 + a_1z^{-1} + \ldots\ldots + a_nz^{-n}}\zeta(t+k) \tag{6.9}$$

Dividing equation 6.9 by $A(z^{-1})$ gives the infinite sequence

$$\begin{aligned}\frac{C}{A}\zeta(t+k) &= [\zeta(t+k) + n_1\zeta(t+k-1) + \ldots\ldots + n_{k-1}\zeta(t+1) \\ &\quad + n_k\zeta(t) + n_{k+1}\zeta(t-1) + \ldots\ldots] \\ &= \tilde{n}(t+k) + \hat{n}(t+k)\end{aligned} \tag{6.10}$$

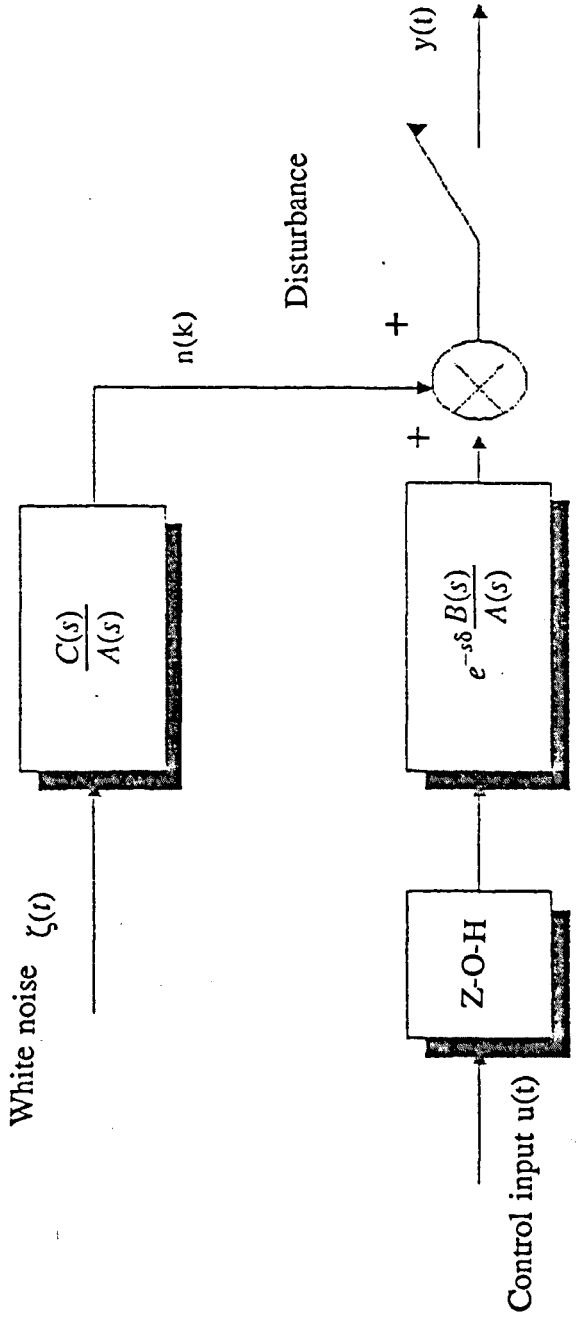


Figure 6.9: Process and noise model

At time t $\tilde{n}(t+k)$ is unknown and cannot be predicted since we have assumed that $\zeta(t)$ is a white noise sequence. The second term $\hat{n}(t+k)$ however, represents the noise signal up to and including time t which can be computed exactly from equation 6.6

$$\zeta(t) = \frac{Ay(t) - z^{-k}Bu(t)}{C} \tag{6.11}$$

Define the identity

$$\frac{C}{A} = E + z^{-k}\frac{F}{A} \tag{6.12}$$

where

$$E[z^{-1}] = 1 + e_1 z^{-1} + ... + e_{k-1} z^{-k+1}$$
$$F[z^{-1}] = f_0 + f_1 z^{-1} + ... + f_{n-1} z^{-n+1}$$

Multiplying equation 6.11 by $\zeta(t+k)$

$$\frac{C}{A}\zeta(t+k) = E\zeta(t+k) + \frac{F}{A}\zeta(t) \tag{6.13}$$

Using equation 6.10

$$\frac{C}{A}\zeta(t+k) = \tilde{n}(t+k) + \hat{n}(t+k) \tag{6.14}$$

The idea is to choose the control u(t) to minimise

$$I = E[y^2(t+k)]$$
$$= E\left[\left(\frac{B}{A}u(t) + \hat{n}(t+k) + \tilde{n}(t+k)\right)^2\right]$$

$$= E\left[\left(\frac{B}{A}u(t) + \hat{n}(t+k)\right)^2 + 2\left(\frac{B}{A}u(t) + \hat{n}(t+k)\right)\tilde{n}(t+k) + \hat{n}^2(t+k)\right] \tag{6.15}$$

Since at time t $\tilde{n}(t+k)$ is unknown and impredictable, it is uncorrelated with both u(t) and $\hat{n}(t+k)$ so that the second term in the above equation goes to zero to leave

$$I = E\left[\left(\frac{B}{A}u(t) + \hat{n}(t+k)\right)^2 + \tilde{n}^2(t+k)\right] \tag{6.16}$$

Differentiating I with respect to u(t)

$$\frac{dI}{du(t)} = 2E\left[\left(\frac{B}{A}u(t) + \hat{n}(t+k)\right)b_0\right] = 0 \tag{6.17}$$

Hence the minimum variance control is

$$u(t) = -\frac{A}{B}\hat{n}(t+k) \tag{6.18}$$

The effect of the control can be seen by substituting equations 6.10 and 6.18 into equation 6.6 to give

$$\begin{aligned} y(t+k) &= \frac{B}{A}\left[-\frac{B}{A}\hat{n}(t+k)\right] + \tilde{n}(t+k) + \hat{n}(t+k) \\ &= \tilde{n}(t+k) \end{aligned} \tag{6.19}$$

Hence the control in equation 6.18 is designed to ensure that the controlled part of the system output cancels the predicted disturbance $\tilde{n}(t+k)$ only. This implies that minimum variance control will be most effective when the prediction is good. To implement the control law, u(t) must be expressed in another form.

From equation 6.14 and using 6.11

$$\begin{aligned} \hat{n}(t+k) &= \frac{F}{A}\zeta(t) \\ &= \frac{F}{A}\left[\frac{Ay(t) - z^{-k}Bu(t)}{C}\right] \end{aligned} \tag{6.20}$$

substituting equation 6.20 into 6.18

$$u(t) = -\frac{A}{B}\frac{F}{A}\left[\frac{Ay(t) - z^{-k}Bu(t)}{C}\right] \tag{6.21}$$

or

$$u(t)\left[\frac{C-z^{-k}}{C}\right] = -\frac{FA}{BC}y(t) \tag{6.22}$$

From equation 6.12

$$C - z^{-k}F = AE \tag{6.23}$$

hence

$$u(t) = -\frac{F}{BE}y(t) \tag{6.24}$$

Equation 6.24 represents the feedback control law that gives minimum variance. The controlled system is illustrated in Figure 6.10

Setting G=BE equation 6.24 becomes

$$Gu(t) = Fy(t) \tag{6.25}$$

Since the plant parameters are unknown, the polynomials $F(z^{-1})$ and $G(z^{-1})$ may be found by estimation. There are two well known methods for this: the implicit and explicit approaches.

6.3.2) The implicit algorithm

For the system whose parameters are unknown, the polynomial $F(z^{-1})$ and $G(z^{-1})$ are estimated on-line . The algorithm relies on implicit rather than explicit identification. The self tuning algorithm employing the implicit identification scheme is shown in Figure 6.11. The process is described by equation 6.5.

$$Ay(t) = z^{-k}Bu(t) + C\zeta(t)$$

multiplying by $E(z^{-1})$ gives

$$AEy(t) = z^{-k}BEu(t) + CE\zeta(t) \tag{6.26}$$

using equation 6.12

$$y(t+k) = \frac{F}{C}y(t-k) + \frac{BE}{C}u(t-k) + E\zeta(t) \tag{6.27}$$

$$y(t+k) = \frac{F}{C}y(t-k) + \frac{G}{C}u(t-k) + \tilde{n}(t+k) \tag{6.28}$$

Substituting equation 6.28 into 6.15 and assuming $C(z^{-1})$ =1

$$\hat{G}u(t) = \hat{F}y(t) \tag{6.29}$$

therefore a recursive least squares algorithm can be used with

$$\left.\begin{aligned} x(t) &= [y(t-k)....y(t-k-n_y);u(t-k)....u(t-k-n_u)]^T \\ \Theta(t) &= [\hat{f}_0\hat{f}_1,....\hat{f}_{ny};\hat{g}_0,....,\hat{g}_{nu}] \end{aligned}\right\} \tag{6.30}$$

If $C(z^{-1})$=1 , the estimates will be unbiased.

6.3.3) The explicit algorithm

The process is described by equation 6.48 and the explicit identification schemes is shown in Figure 6.12.

Define the vectors

$$\left.\begin{aligned} x(t) &= [y(t-1)....y(t-n);u(t-k)....u(t-k-n)]^T \\ \Theta(t) &= [\hat{a}_1,\hat{a}_2,.....,\hat{a}_n;\hat{b}_0,.....,\hat{b}_n] \end{aligned}\right\} \quad (6.31)$$

The polynomial $C(Z^{-1})$ is assumed to be equal to 1 unless an extended recursive least squares estimator is used.

$\Theta(t)$ and x(t) can be used in a recursive least squares estimator to find estimates of the process parameters which may be used to solve equation 6.12 (called diophantine). But equation 6.12 can be difficult to solve.

6.3.4) The generalised minimum variance self-tuning controller

Briefly, consider the system governed by the difference equation

$$y(t) = z^{-k}\frac{B}{A}u(t) + \frac{C}{A}\zeta(t) \quad (6.32)$$

where

$$A = 1 + a_1z^{-1} + ... + a_{na}z^{-na}$$
$$B = b_0 + b_1z^{-1} + ... + b_{nb}z^{-nb}$$
$$C = 1 + C_1z^{-1} + ... + C_{nc}z^{-nc}$$

u(t) and y(t) are the system's input and output sequences, $\zeta(t)$ a zero mean white noise sequence and k is the system's integer time delay. A general minimum variance controller is required to minimum the cost function

$$I_1 = E[(Py(t+k) - Rw(t)^2 + (Q'u(t))^2] \quad (6.33)$$

E being the expectation operator and P, Q and R are discrete polynomials and w(t) is the demand signal .

Put equation 6.32 into its k-step ahead prediction form and multiply both sides by A

$$Ay(t+k) = Bu(t) + C\zeta(t+k) \quad (6.34)$$

Using the identity of equation 6.12

$$E = \frac{C}{A} - z^{-k}\frac{F}{A} \quad (6.35)$$

where the discrete polynomials E and F have degrees (k-1) and (n_a–1) respectively.

Multiplying both sides of equation (6.34) by E

$$EAy(t+k) = EBu(t) + EC\zeta(t+k) \tag{6.36}$$

from equation 6.35

$$EA = C - z^{-k}F \tag{6.37}$$

Substituting for EA in equation 6.35

$$Cy(t+k) = Fy(t) + EBu(t) + EC\zeta(t+k) \tag{6.38}$$

or

$$y(t+k) = \frac{F}{C}y(t) + \frac{EB}{C}u(t) + E\zeta(t+k) \tag{6.39}$$

Substituting for y(t+k) in the cost function I_1 , yields

$$I_1 = E[(\frac{PF}{C}y(t) + \frac{PEB}{C}u(t) + PE\zeta(t+k) - Rw(t))^2 + (Q'u(t))^2] \tag{6.40}$$

Provided that the order of P is less than k, $Pe\zeta(t+k)$ is uncorrelated with the rest of the quantities inside the expectation operator and I_1 can be re-written as

$$I_1 = E\left[(\frac{PF}{C}y(t) + \frac{PEB}{C}u(t) - Rw(t))^2 + (Q'u(t))^2\right] + E\left[(PE\zeta(t+k))^2\right] \tag{6.41}$$

Differentiating I_1 with respect to u(t) and equating the result to zero

$$\frac{dI_1}{du(t)} = 2b_0\left[(\frac{PF}{C}y(t) + \frac{PEB}{C}u(t) - Rw(t)\right] + 2q'_0Q'u(t) = 0 \tag{6.42}$$

where

$e_0 = p_0 = 1$ is assumed without loss of generality.

Now let

$$Q = q'_0\frac{Q'}{b_0}$$

to give

$$PFy(t) + (PEB + CQ)u(t) - RCw(t) = 0 \tag{6.43}$$

or

$$\bar{F}y(t) + \bar{G}u(t) + \bar{H}w(t) = 0 \tag{6.44}$$

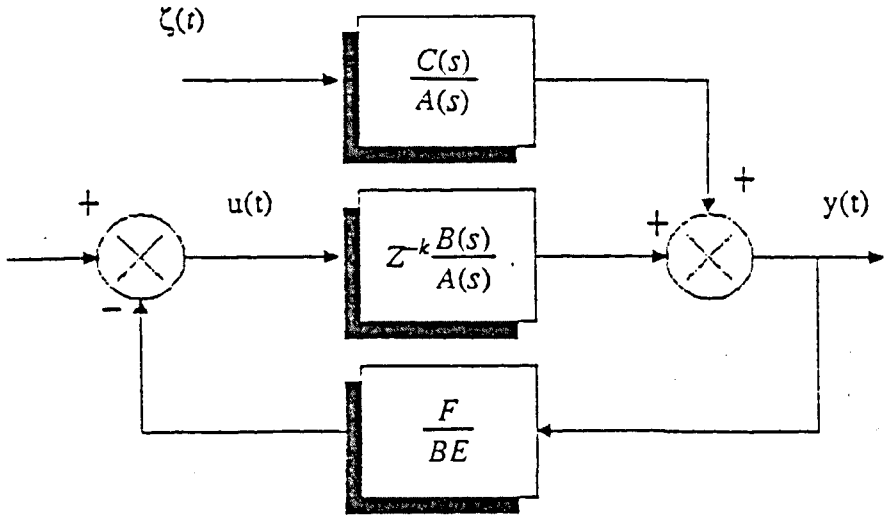


Figure 6.10: Block diagram of the controlled system

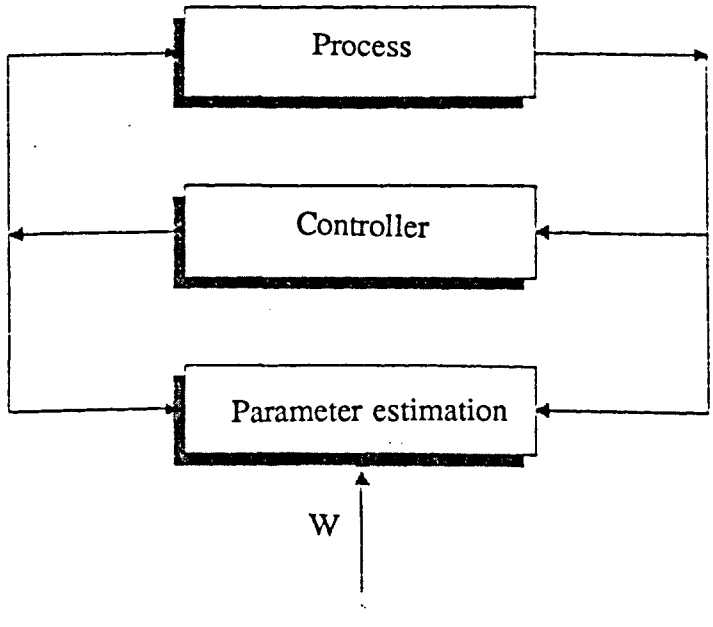


Figure 6.11: Implicit identification

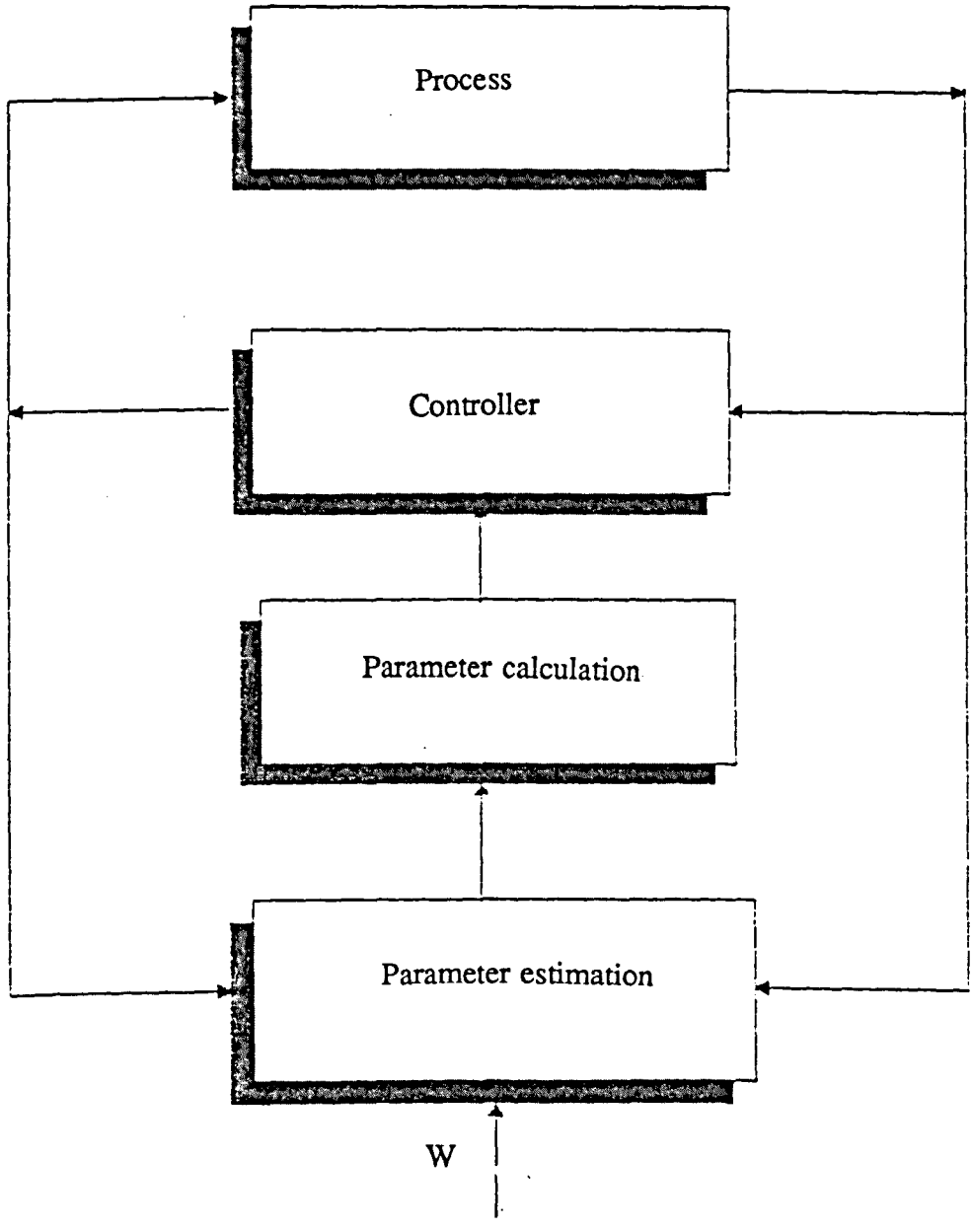


Figure 6.12: Explicit identification

where

$$\overline{F} = PF$$
$$\overline{G} = PG + CQ = PBE + CQ$$
$$\overline{H} = -RC$$

Alternatively, if the concept of the generalised (auxiliary) output function of the form

$$\Phi(t) = Py(t) - Rw(t-k) + Qu(t-k) \tag{6.45}$$

or

$$\Phi(t+k) = Py(t+k) - Rw(t) + Qu(t) \tag{6.46}$$

(k-step ahead predictor form) is considered and this is shown in Figure 6.13.

Now consider a control law that seeks to minimise the variance of this auxiliary output

$$I_2 = E\left[(\Phi(t+k))^2\right] \tag{6.47}$$

$$I_2 = E\left[\left(P\frac{F}{C}y(t) + \frac{BE}{C}u(t) + \tilde{n}(t+k) - Rw(t) + Qu(t)\right)^2\right] \tag{6.48}$$

$$\frac{dI_2}{du(t)} = P\left[\frac{F}{C}Z(t) + \frac{BE}{C}u(t)\right] - Rw(t) + Qu(t))2(b_0 + q_0) = 0 \tag{6.49}$$

and hence the control law

$$F'y(t) + G'u(t) + H'w(t) = 0 \tag{6.50}$$

where

$$F' = PF$$
$$G' = PG + CQ$$
$$H' = -CR$$

as previously

Thus, minimisation of the two cost functions I_1 and I_2 achieve the same control law and minimise the variance of the auxiliary output. If the parameters are unknown they will need to be estimated. From equations 6.40, 6.44 and 6.46

$$\begin{aligned}\Phi(t+k) &= \overline{F}y(t) + \overline{G}u(t) + \overline{H}w(t) + \tilde{n}(t+k) \\ &= X^T(t)\Theta + \tilde{n}(t+k)\end{aligned} \tag{6.51}$$

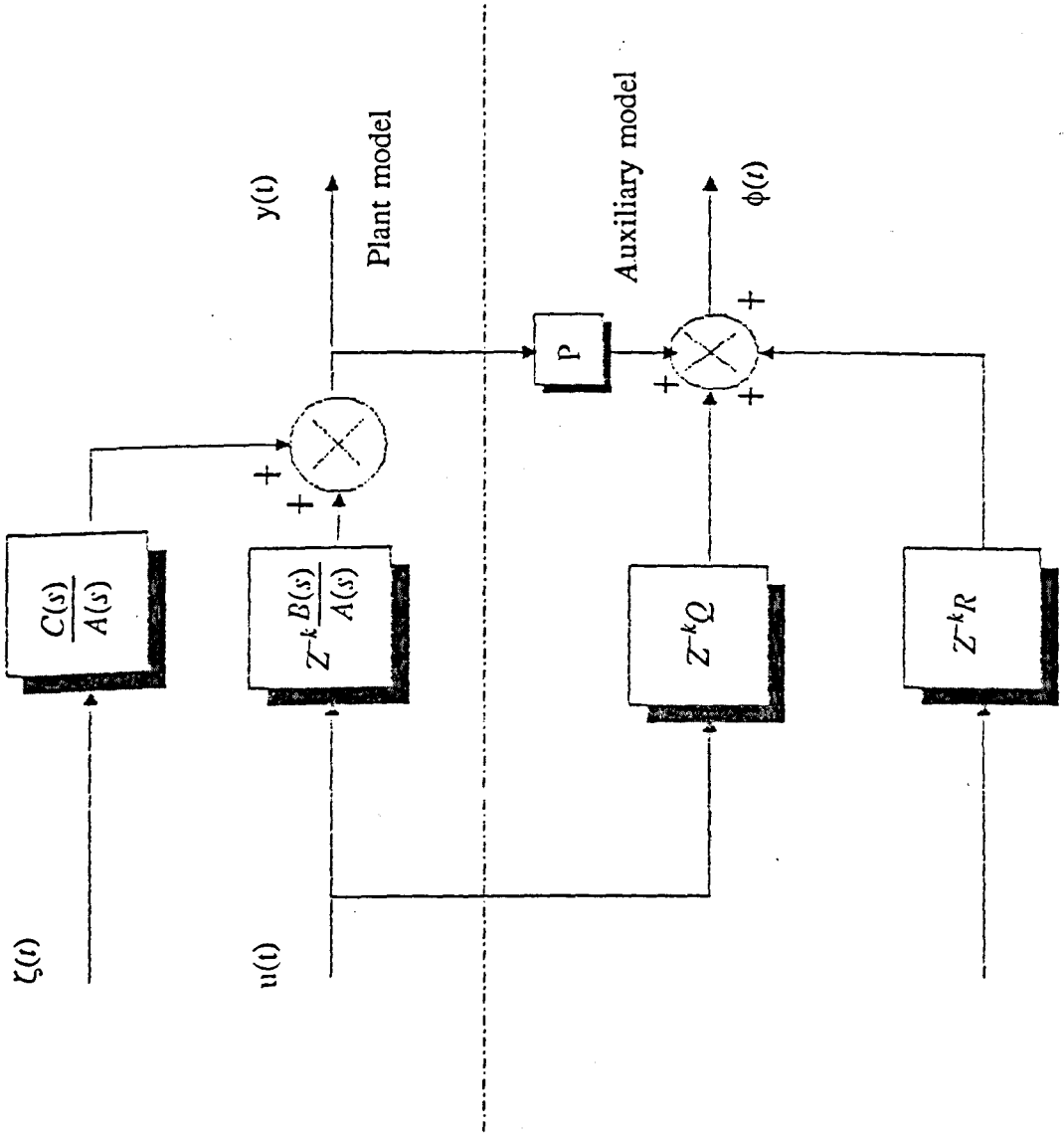


Figure 6.13: Repesentation of the auxiliary output

where

$$X^T(t) = [y(t-k),y(t-k-1),....;u(t-k),....;w(t-k),....]$$
$$\Theta^T = [\overline{f}_0\overline{f}_1,....;\overline{g}_0,....;\overline{h}_0,....] \quad (6.52)$$

The parameters of Θ^T can be estimated using either an ordinary or extended recursive least squares estimator.

The general minimum variance algorithm described previously is applied to control the NARMAX models of both Vecuronium and Atracurium. Eventhough the algorithm has been applied largely to linear models, in the next section of this chapter an attempt with the NARMAX models will be shown.

6.3.5) GMV conrol of muscle relaxation

In what follows, the above developed GMV controller is employed in the closed loop control of muscle relaxation. Here again, both the Vecuronium and Atracurium models are considered with the latter being represented by either of equations 4.5, 4.9 and 4.10.

a) Case of the Vecuronium model

The simulation for this case was run with the following specifications

Order of P-polynomial NP=0 with $p_0 = 1.0$

Order of Q-polynomial NQ=0 with $q_0 = 0.2$

Order of R-polynomial NR=0 with $r_0 = 1.0$

The system and the noise polynomial orders where set to:

$N_a = 5$, $N_b = 4$ and $N_c = 0$ respectively, while the forgetting factor and the initial diagonal elements of the covariance matrix to: 0.995 and 1000 respectively.

The system dead-time for this case assumed a value of 3. The initial vector of the controller parameter estimates was

$$[0,0,0,0,0;1,0,0,0,0;-1,0,0,0,0]$$
$$\overline{f}; \qquad \overline{g}; \qquad \overline{h}$$

The reference (demand) signal was a square wave of levels 0.8-0.4 and 0.8-0.6.

The output and control signals for this case are shown in Figures 6.14 and 6.15.

b) Case of the Atracurium model (NLI)

With the exception of N_a , N_b and q_0 and the system time delay which now assume values of 6, 4, 0.12 and 4 respectively, the simulation was run with the same specifications as those for case a) above. The resulting output and control signals are shown in Figures 6.16 and 6.17.

c) Case of the Atracurium model (NOI)

Figures 6.18 and 6.19 show the system responses for this case. the simulation conditions are identical to those of the NLI-model except for q_0 which is now 0.22.

6.4) Generalised predictive control (GPC) algorithm.

The robustness of the existing self-tuners is open to question when assumptions have to be made about the model order and/or the dead time. Minimum variance (Astrom and Wittenmark, 1973) and generalised minimum variance (Clarke and Gawthrop, 1975, 1979) self-tuning controllers (or regulators) are particularly sensitive to varying time delays. Self-tuning controllers based on the pole assignment principle (Wellstead et al, 1979), though capable of coping with variable delays, may perform unacceptably if the system order is over estimated. The GPC algorithm (Clarke et al, 1987) appears to overcome both of these above problems. The ability of this algorithm to cope with varying or unknown dead times follows form its 'explicitness' while its predictive nature makes it suitable for handling the over parametrisation problem. The remaining of this chapter introduces the basic GPC algorithm.

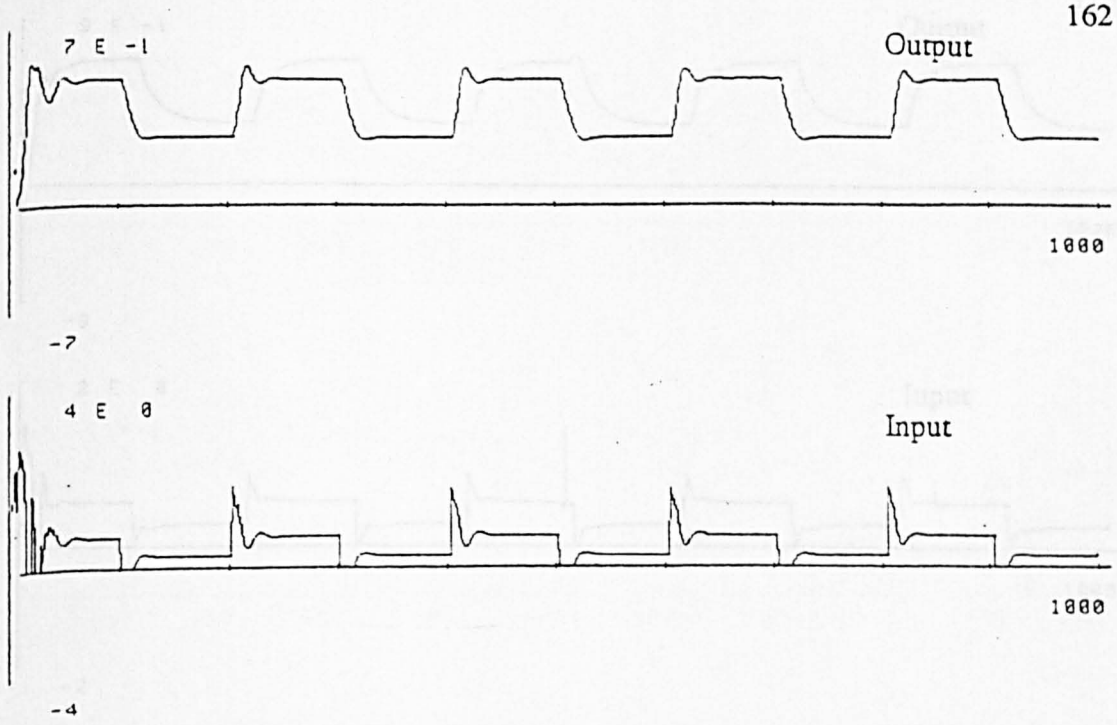


Figure 6.14: Control of the Vecuronium model (NLI) using GMV

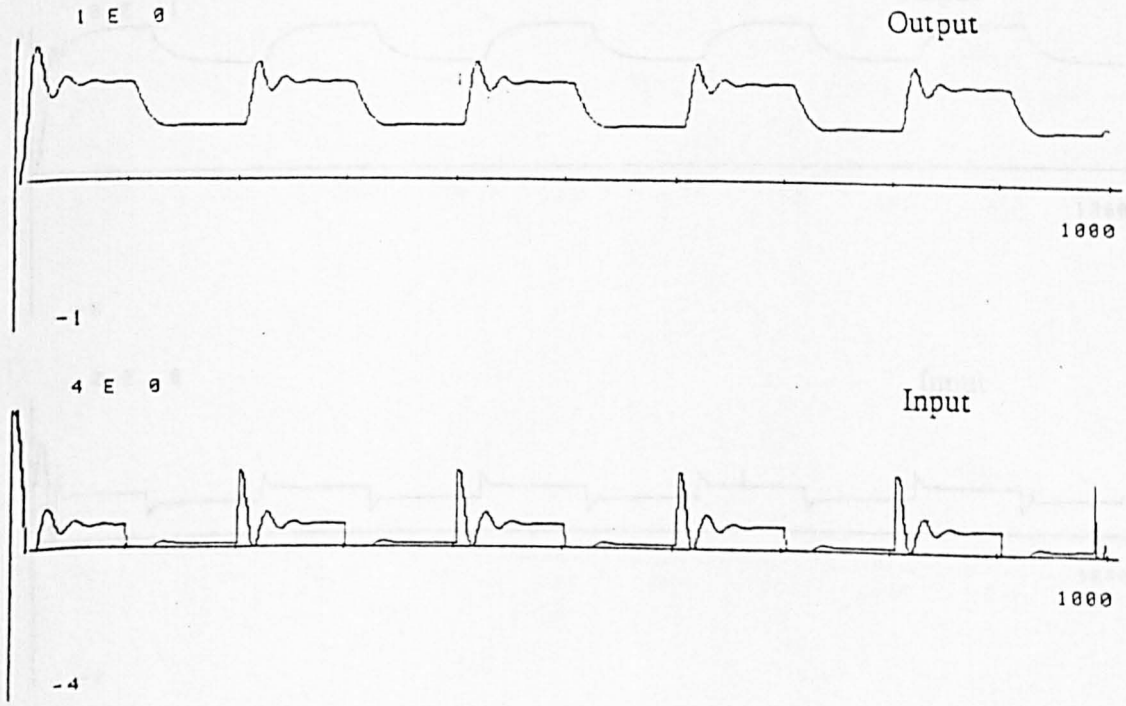


Figure 6.15: Control of the Vecuronium model (NLI) using GMV

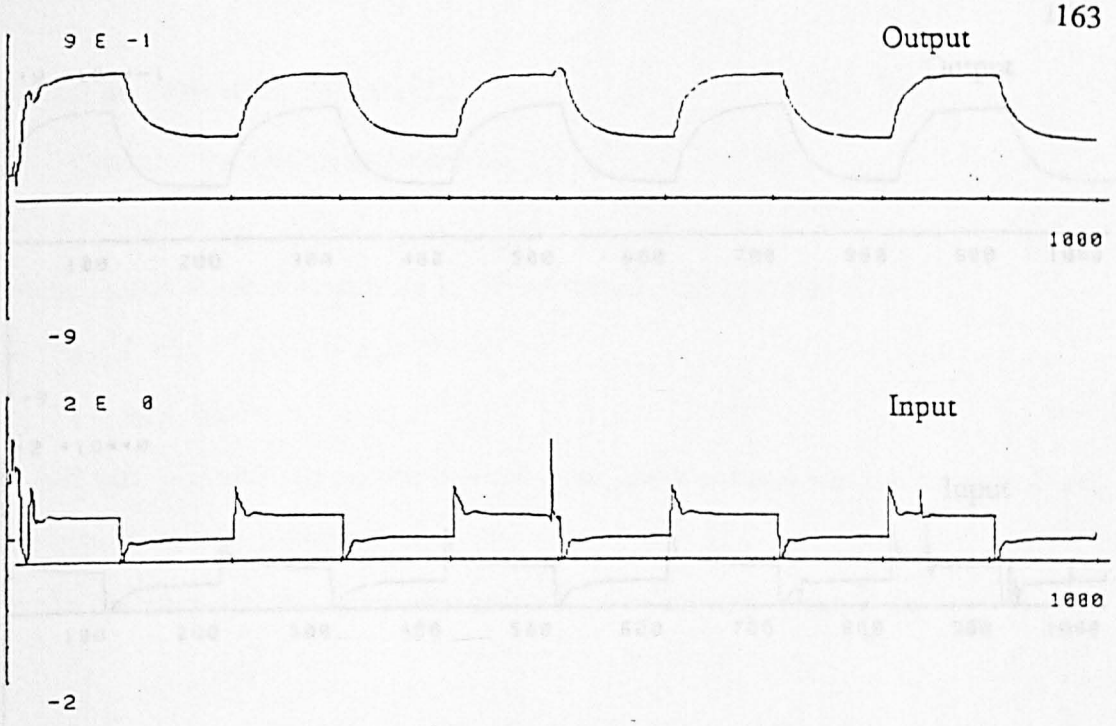


Figure 6.16: Control of the Atracurium model (NLI) using GMV

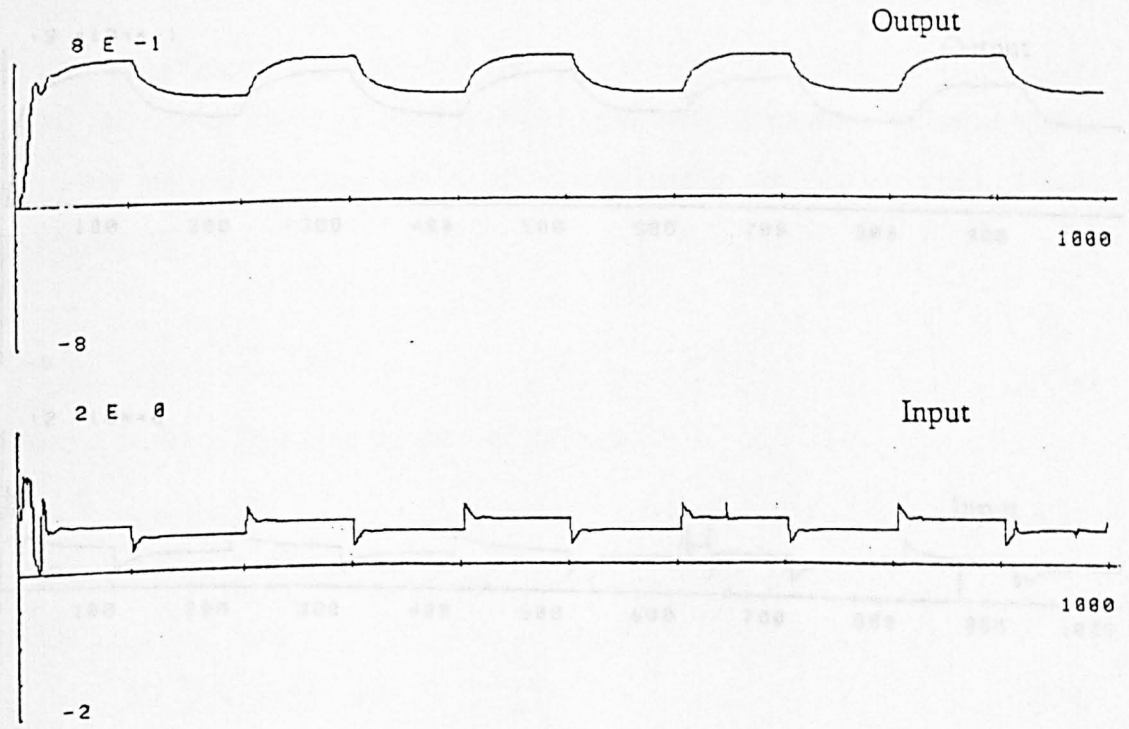


Figure 6.17: Control of the Atracurium model (NLI) using GMV

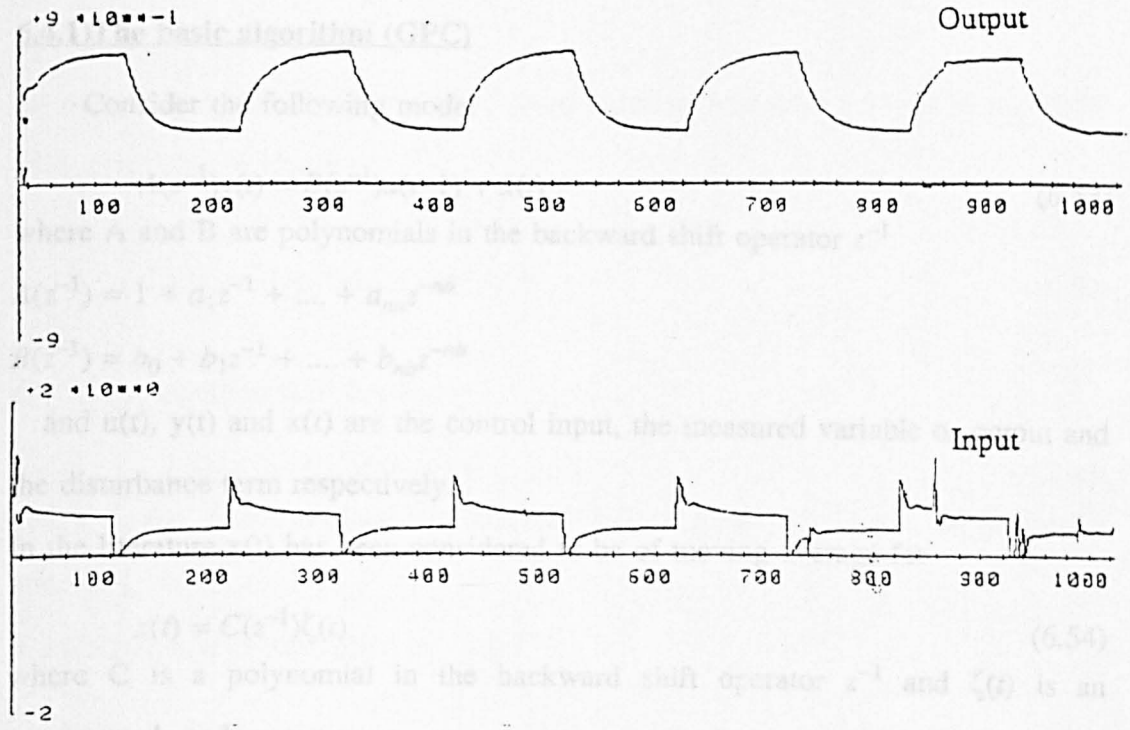


Figure 6.18: Control of the Atracurium model (NOI) using GMV

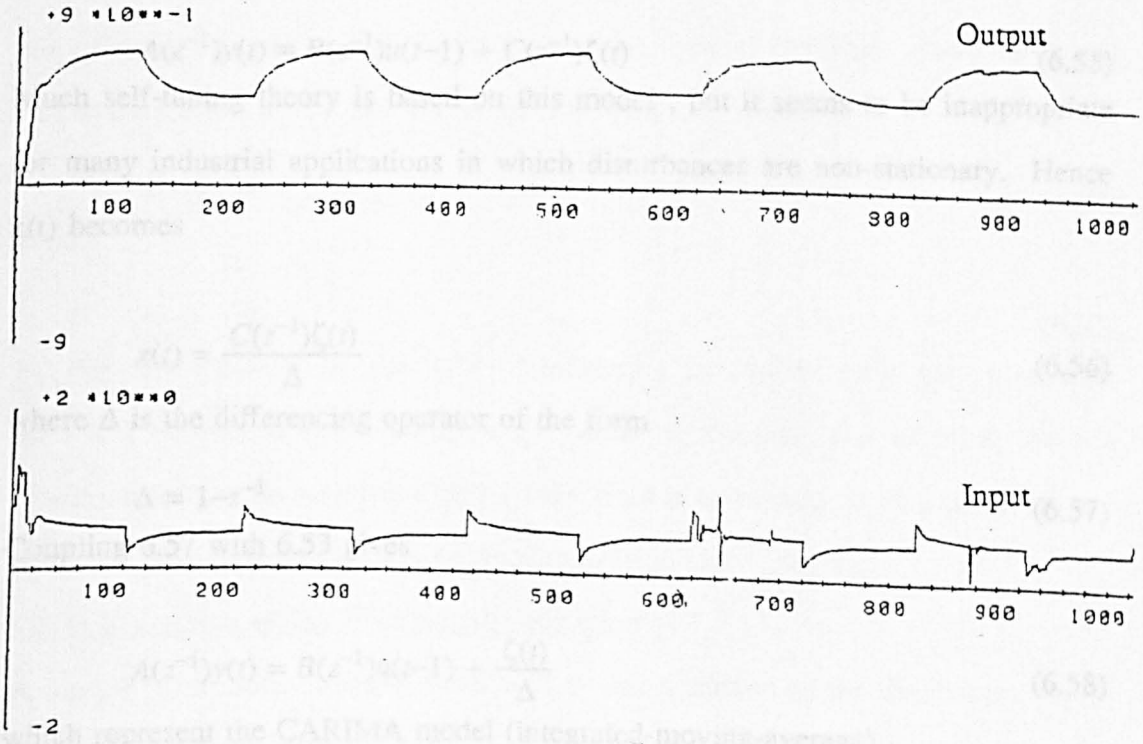


Figure 6.19: Control of the Atracurium model (NOI) using GMV

6.4.1)The basic algorithm (GPC)

Consider the following model

$$A(z^{-1})y(t) = B(z^{-1})u(t-1) + x(t) \tag{6.53}$$

where A and B are polynomials in the backward shift operator z^{-1}

$A(z^{-1}) = 1 + a_1 z^{-1} + ... + a_{na} z^{-na}$

$B(z^{-1}) = b_0 + b_1 z^{-1} + ... + b_{nb} z^{-nb}$

and u(t), y(t) and x(t) are the control input, the measured variable or output and the disturbance term respectively.

In the literature x(t) has been considered to be of moving average for

$$x(t) = C(z^{-1})\zeta(t) \tag{6.54}$$

where C is a polynomial in the backward shift operator z^{-1} and $\zeta(t)$ is an uncorrected random sequence

By combining equation 6.54 with 6.53 a CARMA (Controlled Auto-Regressive and Moving-Average) model is obtained.

$$A(z^{-1})y(t) = B(z^{-1})u(t-1) + C(z^{-1})\zeta(t) \tag{6.55}$$

Much self-tuning theory is based on this model , but it seems to be inappropriate for many industrial applications in which disturbances are non-stationary. Hence x(t) becomes

$$x(t) = \frac{C(z^{-1})\zeta(t)}{\Delta} \tag{6.56}$$

where Δ is the differencing operator of the form

$$\Delta = 1 - z^{-1} \tag{6.57}$$

Coupling 6.57 with 6.53 gives

$$A(z^{-1})y(t) = B(z^{-1})u(t-1) + \frac{\zeta(t)}{\Delta} \tag{6.58}$$

which represent the CARIMA model (integrated-moving-average)

with $C(z^{-1})$ =1 for simplicity

To derive the j-step ahead predictor y(t+j) based on equation 6.58, as in the GMV case, consider the identity

$$C(z^{-1}) = E_j(z^{-1})A\Delta + z^{-j}F_j(z^{-1}) \tag{6.59}$$

with

$C(z^{-1})$ =1, equation 6.59 becomes

$$1 = E_j(z^{-1})A\Delta + z^{-j}F_j(z^{-1}) \tag{6.60}$$

where E_j and F_j are polynomials uniquely defined given $A(z^{-1})$ and the prediction interval j.

If equation 6.58 is multiplied by $E_j\Delta z^j$

$$E_jA\Delta y(t+j) = EB\Delta u(t+j-1) + E_j\zeta(t+j)$$

and substituting for $E_jA\Delta$ from equation 6.60

$$y(t+j) = E_jB\Delta u(t+j-1) + F_jy(t) + E_j\zeta(t+j) \tag{6.61}$$

where y(t+j) is the j-step ahead predictor. As $E_j(z^{-1})$ is of degree j-1 the noise components are well in the future so that the optimal predictor, given measured output data up to time t and any given u(t+i) for i>1, is clearly

$$\hat{y}(t+j/t) = G_j\Delta u(t+j-1) + F_jy(t) \tag{6.62}$$

where

$$G_j = E_jB \tag{6.63}$$

In the development of the GMV self-tuning controller only one prediction $\hat{y}(t+k/t)$ is used where k assumed the value of the plant's dead time. Here a whole set of predictions for which j runs from a minimum up to a large value : these are called the minimum and maximum "prediction horizon".

6.4.2) Recursion of the Diophantine equation

A simple and more effective scheme is to use recursion of the diophantine equation so that the polynomials E_{j+1} and F_{j+1} are obtained given the values of E_j and

F_j.

For clarity of notation set

$$\tilde{A} = A\Delta \ , \ E=E_j \ , \ F=F_j \ , \ R=E_{j+1} \ , \ S=F_{j+1}$$

equation 6.60 gives

$$1 = E\tilde{A} + z^{-j}F \tag{6.64}$$

$$1 = R\tilde{A} + z^{-(j+1)}S \tag{6.65}$$

This is a recursion of the Diophantine equation relating E_j to E_{j+1} and F_j to F_{j+1}

Substracting equation 6.64 from 6.65

$$\tilde{A}(R-E) + z^{-j}(z^{-1}S-F) = 0 \tag{6.66}$$

The polynomial E_j is of degree j-1, hence R is of degree j so is R-E and therefore

$$R-E = \tilde{R} + r_j z^{-j} \tag{6.67}$$

Equation 6.66 becomes

$$\tilde{A}(\tilde{R} + r_j z^{-j}) + z^{-j}(z^{-1}S - F) = 0 \tag{6.68}$$

Hence

$$\tilde{A}\tilde{R} + z^{-j}(z^{-1}S - F + \tilde{A}r_j) = 0 \tag{6.69}$$

For equation 6.69 to be equal to zero

$$\begin{Bmatrix} \tilde{R} = 0 \\ z^{-1}S - F + \tilde{A}r_j = 0 \end{Bmatrix} \ OR \ \begin{Bmatrix} \tilde{R} = 0 \\ S = z(F - \tilde{A}r_j) \end{Bmatrix} \tag{6.70}$$

Let us consider the following equation

$$\begin{cases} S = z(F_j - \tilde{A}r_j) \\ \tilde{A} = A\Delta \end{cases} \tag{6.71}$$

Where

$$A = 1 + a_0 z^{-1} + \ldots\ldots + a_{na} z^{-na}$$

$$F = f_0 + f_1 z^{-1} + \ldots\ldots + f_{nf} z^{-nf}$$

As 1 is a leading element of A so $\tilde{A}$ and f_0 first element of F_j , then $r_j = f_0$ because S is of power z^{-j} and $(F_j - \tilde{A}r_j)$ is multiplied by z hence

$$r_j = f_0 \tag{6.72}$$

which is the first element of the polynomial F_j , and the following components of

higher power are obtained using the recursion

$$S_i = f_{i+1} - \tilde{a}_{i+1} r_j \tag{6.73}$$

where

S_i is the next horizon

f_{i+1} and r_j are the last horizon.

Setting $\tilde{R} = 0$ equation 6.67 becomes

$$R = E + r_j z^{-j} \tag{6.74}$$

$$E_{j+1} = E_j + r_j z^{-j} \tag{6.75}$$

and using equation 6.63

$$G_{j+1} = B(z^{-1})R(z^{-1}) \tag{6.76}$$

So equations 6.72 to 6.76 constitute the basis for the implementation of the GPC algorithm.

6.4.3) The predictive control law

Suppose a future set-point or reference sequence of the form

$$w(t+j) \;\; ; \;\; j = 1,2,..... \tag{6.77}$$

is available. In the GPC case a smoothed approach from the current output y(t) to w(t) is required which is obtainable from the simple first order lag model

$$w(t) = y(t)$$

$$w(t+j) = \alpha w(t+j-1) + (1-\alpha)w$$

where $\alpha = 1$ for a slow transition from the current measured variable to the real set point. This approach is illustrated in Figure 6.20

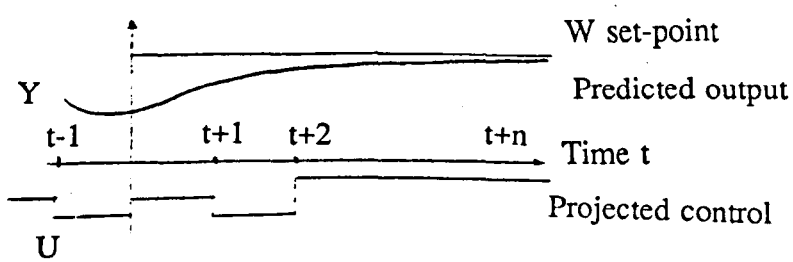


Figure 6.20: Set-point control and output in GPC

The aim of the predictive control law is to drive future plant outputs y(t+j) close

w(t+j).

Consider the following cost function

$$I = I_G(e,\tilde{u}) \tag{6.78}$$

where

$\tilde{u}$ is a vector of increments of u and I is a quadratic function of the form

$$I_G = \sum_{j=N_1}^{j=N_2} e^2(t+j) + \sum_{j=1}^{j=N_u} \lambda \Delta u^2(t+j-1) \tag{6.79}$$

The method could be summarized as follows:

At each present time t, a forecast is made of the process output over a long range time horizon by means of a mathematical model of the process dynamics and is a function of the future control weighted to apply at time t. As a result of this forecast several control actions will be proposed but the best will be selected so that the predicted output will follow the predicted set-point. The control action is applied at the present time t. The whole procedure is then repeated at the next sample and this is called the receding horizon approach (De Keyser and Vancomwenberghs, 1982).

(1) at sample instant t the response of the plant is

$$[y(t),y(t-1);u(t-1)] \tag{6.80}$$

(2) The control increment vector is computed using the optimization routine

(3) u(t) is extracted and applied to the plant.

(4) all sequences are shifted in preparation for the next sample to repeat the same procedure. The prediction sequence is shown in Figure 6.21.

Consider the cost function repesented by equation 6.79

$$e^2(t+j) = [y(t+j) - w(t+j)]^2 \tag{6.81}$$

N_1 is the minimum costing horizon

N_2 is the maximum costing horizon

λ_j are the control weighting sequences

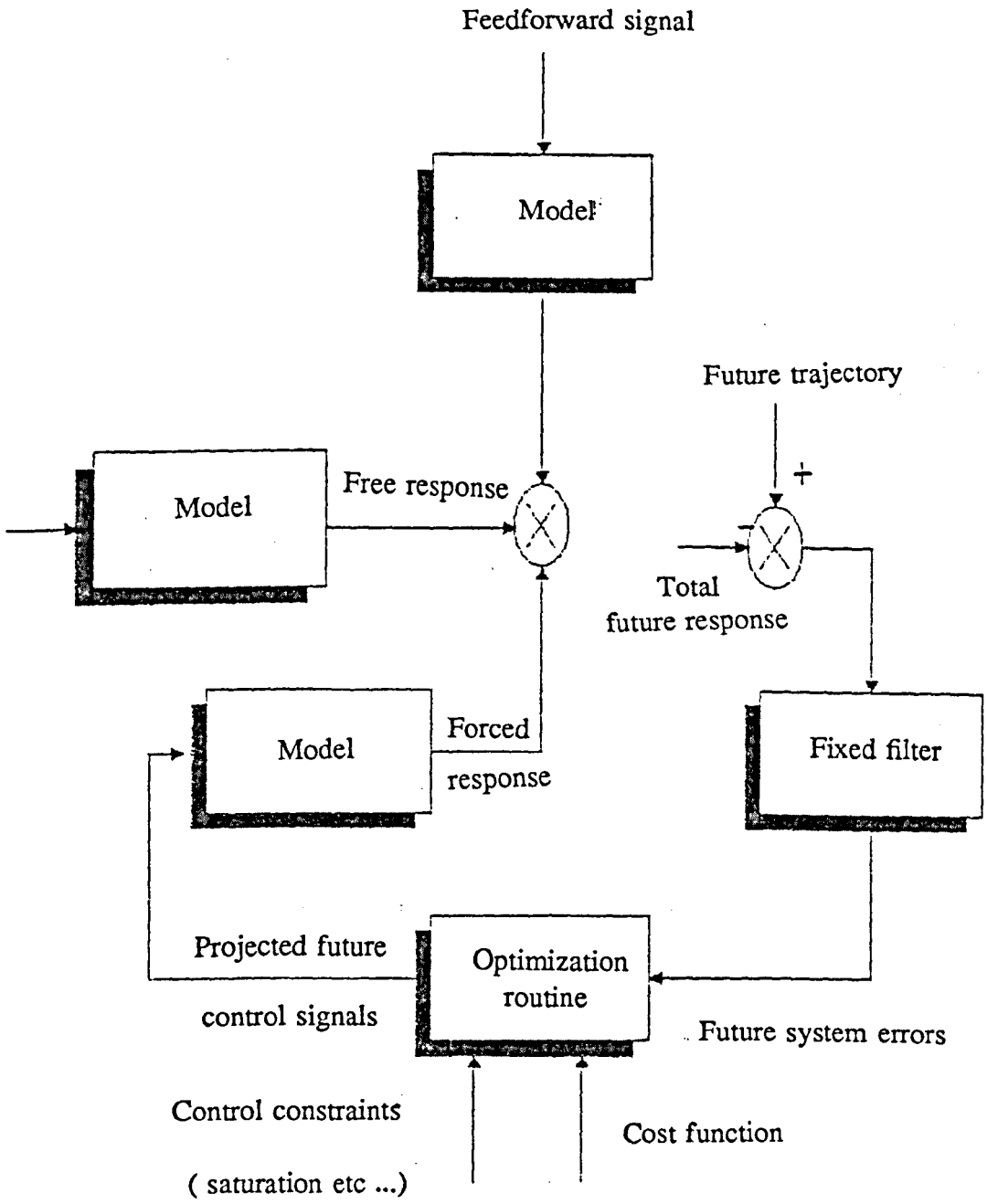


Figure 6.21: The prediction sequence in GPC

In general N_2 is chosen to include all responses which are affected by the current control. It should be greater than the degree of $B(z^{-1})$, it is however set to approximate the rise time plant N_1 can often be considered equal to 1 but the value of the dead-time of the process is known prior equal to "K".

By setting

$$\left\{\begin{matrix} N_1 = 1 \\ N_2 = N \\ \lambda(j) = \lambda \end{matrix}\right\} \tag{6.82}$$

equation 6.61 becomes

$$\left\{\begin{matrix} y(t+1) = G_1\Delta u(t) + F_1 y(t) + E_1\zeta(t+1) \\ y(t+2) = G_2\Delta u(t+1) + F_2 y(t) + E_2\zeta(t+2) \\ . \\ . \\ . \\ . \\ y(t+N) = G_N\Delta u(t+N-1) + F_N y(t) + E_N\zeta(t+N) \end{matrix}\right\} \tag{6.83}$$

Consider the component of y(t+j) which consists of signals known at time t f(t+j) then

$$\left\{\begin{matrix} f(t+1) = [G_1(z^{-1}) - g_{1o}]\Delta u(t) + F_1 y(t) \\ f(t+2) = z(G_2(z^{-1}) - g_{2o} - z^{-1}g_{21})\Delta u(t) + F_2 y(t) \\ . \\ . \\ . \\ .etc \end{matrix}\right\} \tag{6.84}$$

where

$$G_i(z^{-1}) = g_{i0} + g_{i1}z^{-1} + \ldots\ldots\ldots \tag{6.85}$$

So equation 6.83 can be written in the key vector form

$$\hat{y} = G\tilde{u} + f \tag{6.86}$$

where the vectors are all $N\times 1$

$$\hat{y} = [\hat{y}(t+1),\hat{y}(t+2),\ldots\ldots,\hat{y}(t+N)]^T$$

$$\tilde{u} = [\Delta u(t),\Delta u(t+1),.....,\Delta u(t+N-1)]^T$$
$$f = [f(t+1),f(t+2),.....f(t+N)]^T$$

In the $G_i(z^{-1})$ polynomial

k goes from 0 to the present horizon

i the value of the present horizon

$$g_{ik} = g_k, \text{ for } all\ k<i \tag{6.87}$$

The matrix G of equation 6.86 is the lower triangular of dimension $N \times N$ of the form

$$G = \begin{bmatrix} g_{10} & & \\ g_{21} & g_{20} & \\ . & & \\ . & & \\ . & & \\ . & & \\ . & & \\ g_{N,N-1} & & g_{N0} \end{bmatrix} \tag{6.88}$$

Using equation 6.87, equation 6.88 becomes

$$G = \begin{bmatrix} g_0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 & & \\ g_1 & g_0 & \\ g_2 & g_1 & g_0 \\ . & & \\ . & & \\ . & & \\ . & & \\ . & & \\ . & & \\ g_{N-1} & & g_0 \end{bmatrix} \tag{6.89}$$

If the process has a dead-time $K>=1$, the first row of G is zero, if K=1 ; $g_0 = 0$ etc.

Consider the cost function of equation 6.79

$$I_G = \sum_{N_1}^{N_2}(\tilde{y}-W)^2 + \sum_{j=1}^{N_2}\lambda(j)[\Delta u(t+j-1)]^2 \tag{6.90}$$

Minimizing equation 6.90 leads to $\tilde{y} = G\tilde{u} + f$

$$E[I_G] = I_1 = (G\tilde{u} + f - W)^T(G\tilde{u} + f - W) + \lambda\tilde{u}^T\tilde{u} \tag{6.91}$$

$$\tilde{u} = (G^TG + \lambda I)^{-1}G^T(W-f) \tag{6.92}$$

$$\tilde{u} = [\Delta u(t),\Delta u(t+1),.....,\Delta u(t+N-1)]^T \tag{6.93}$$

$\Delta u(t+1)$,.... etc could be ignored and as $\Delta u(t)$ is the first element of $\tilde{u}$, therefore

$$u(t) = u(t-1) + g^T(W-f) \tag{6.94}$$

where

g^T is the first row of $(G^TG + \lambda I)^{-1}G^T$

6.4.4) The control law

Consider the vector e composed of predicted future system error $(w(t+j) - \hat{y}(t+j)$. Using the cost function of equation 6.79

$$I_G(N_1,N_2,N_u,\lambda) = \sum_{j=N_1}^{j=N_2} e^2(t+j) + \lambda\sum_{j=1}^{N_u}\Delta u^2(t+j-1) \tag{6.95}$$

(1) Given y(t) and the past values of y, u, Δu the prediction of the freely responding plant are made to form, by comparison with w, the vector e.

(2) for a given N_1, N_2, N_u, λ , e the vector $\tilde{u}$ is calculated by minimizing I

(3) The first element $u(t) = u(t-1) + \Delta u(t)$ is asserted and sequences are shifted ready for repeating sequence 1, 2, ...

(4) setting N_1 to a value less than the dead-time K, would add unnecessary calculations. If the dead-time is unknown, N_1 is set to 1 and the degree of $B(z^{-1})$ increased

(5) N_2 should exceed the degree of $B(z^{-1})$.

(6) N_u , which is the control horizon determines the degree of freedom in the future control increments. Generally N_u=1 gives satisfactory performances.

(7) λ which the control weighting factor, for the easiest chose is set to zero, but very often to a very small value.

6.4.5) General predictive control of muscle relaxants

As for the GMV approach of section 6.3 the drug dynamics are represented by the identified NARMAX models of chapter 5

a) Case of the drug Vecuronium

The simulation was run over 1000 samples with the following specifications

$N_1 = 1, \quad N_2 = 10, \quad N_u = 1, \quad \lambda = 0.1$.

The orders of the A and B polynomials were set to $n_a = 5$ and $n_b = 6$, and the forgetting factor and the initial diagonal elements of the covariance matrix in the estimation routine to 0.995 and 1000 respectively.

Initial control was via a PI controller whose parameters were set to

$K_p = 0.8$, and $K_e = 0.04$.

Lower and upper limits of 0 and 3 respectively were imposed on the control signal. The output responses and control signals for square wave demands of levels (0.8 and 0.4) and (0.8 and 0.6) are shown in Figures 6.22 and 6.23 respectively. An other two simulation runs were performed with the same specification with the exception of: First the lower and upper limit were set to 0 and 1 respectively (Figure 6.23a) and second with the same limit but the control was via an optimized PID controller were $K_p = 0.95$, $K_e = 0.1$ and $K_i = 2.8$ (Figure 6.23b).

b) Case of the NLI Atracurium model

With the exception of the A and B polynomial orders and the tuning factor λ , which were set to

$n_a = 4$, $n_b = 8$ and $\lambda = 0$ and the control limits of 0 and 1, the simulation for this case was run with the specifications of case a) above. The resulting output and control signals are shown in Figures 6.24 and 6.25.

c) Case of the NOI Atracurium model

Figures 6.26 and 6.27 depict the output and control signals for this case with the simulation run with specifications identical to those in case b) above.

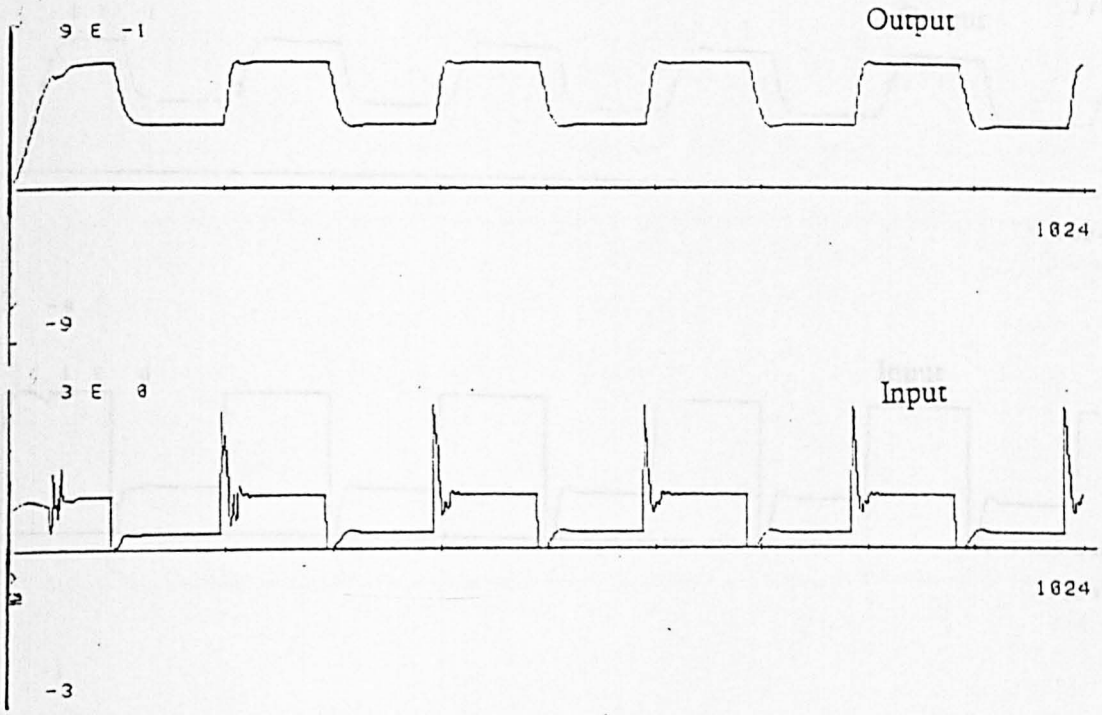


Figure 6.22: Control of the Vecuronium model (NLI) using GPC

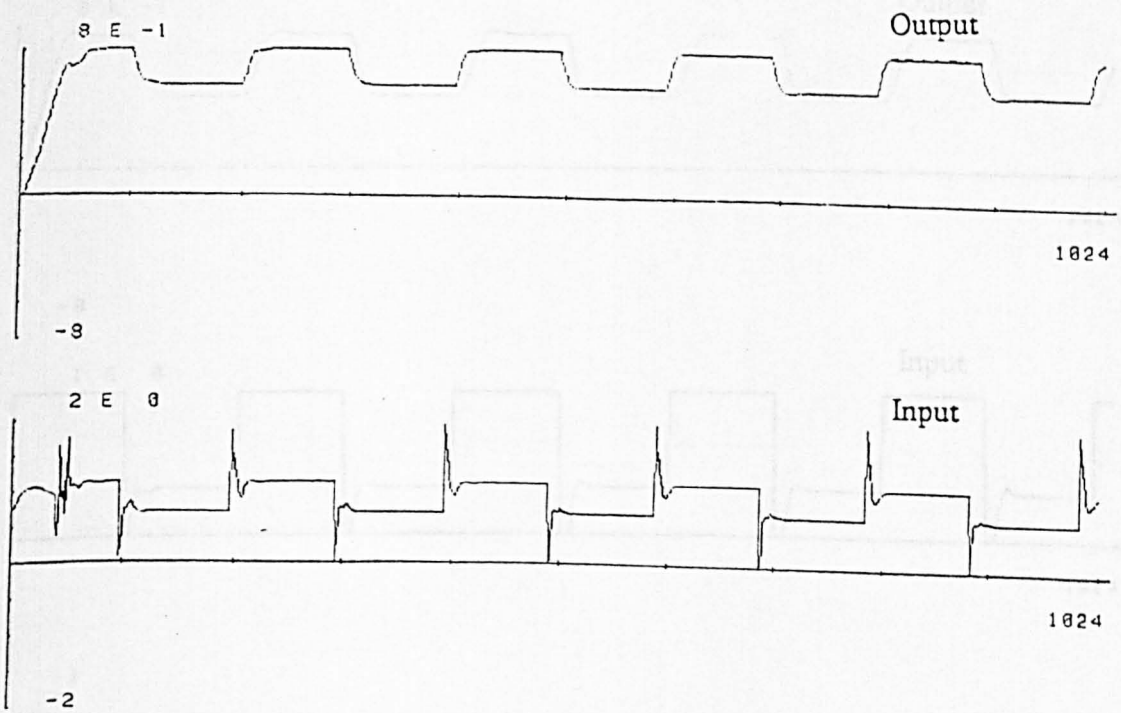


Figure 6.23: Control of the Vecuronium model (NLI) using GPC

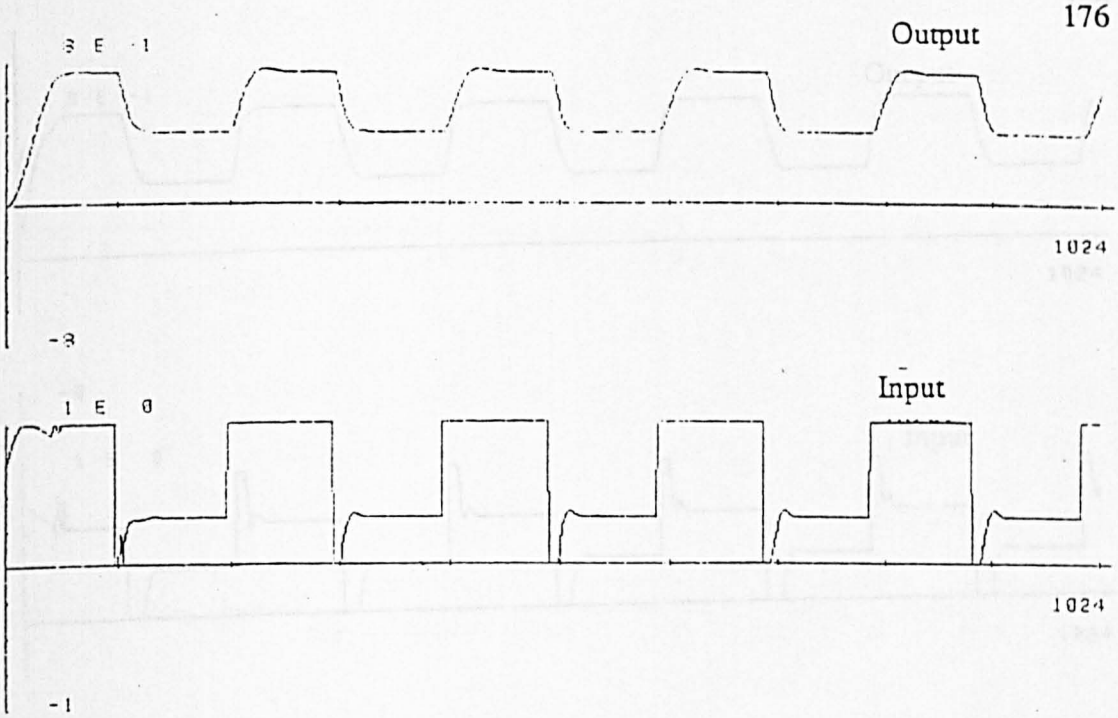


Figure 6.23a: Control of the Vecuronium model (NLI) using GPC

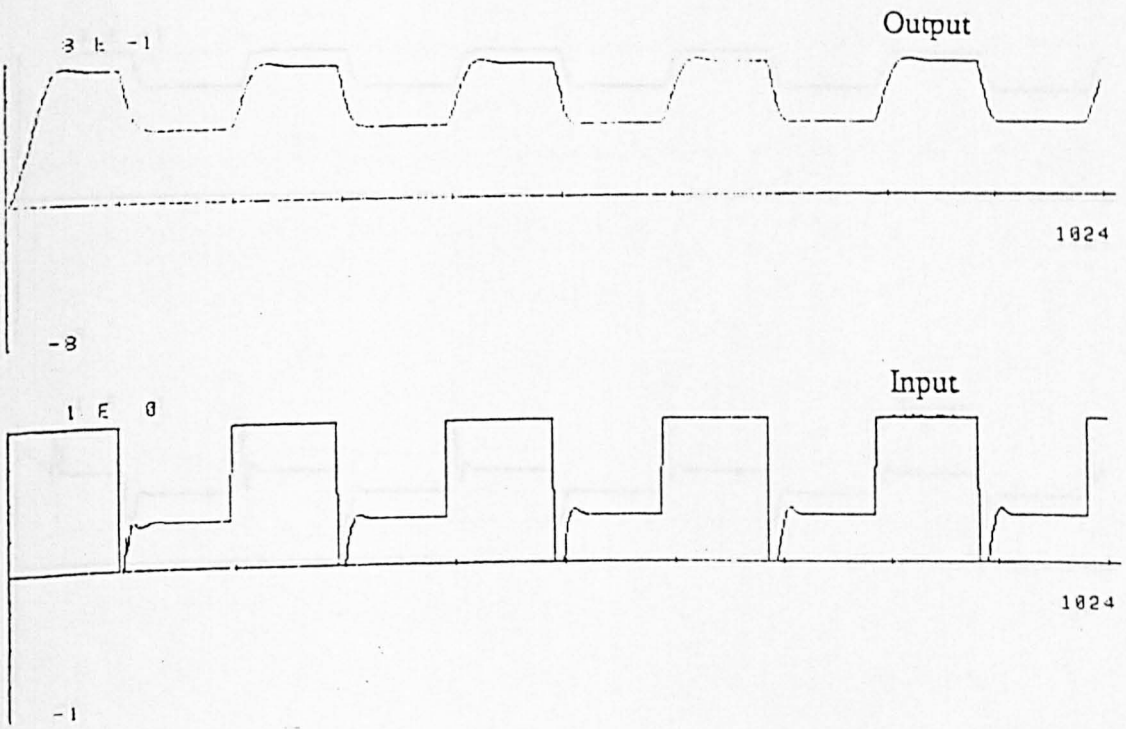


Figure 6.23b: Control of the Vecuronium model (NLI) using GPC

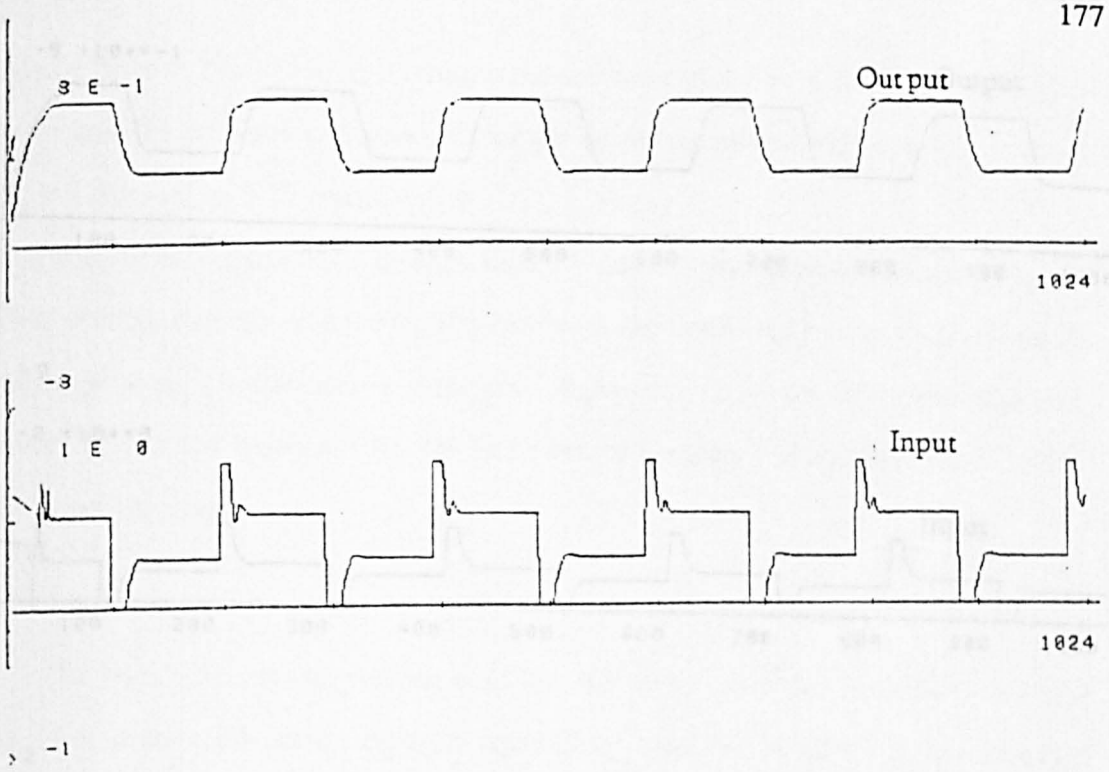


Figure 6.24: Control of the Atracurium model (NLI) using GPC

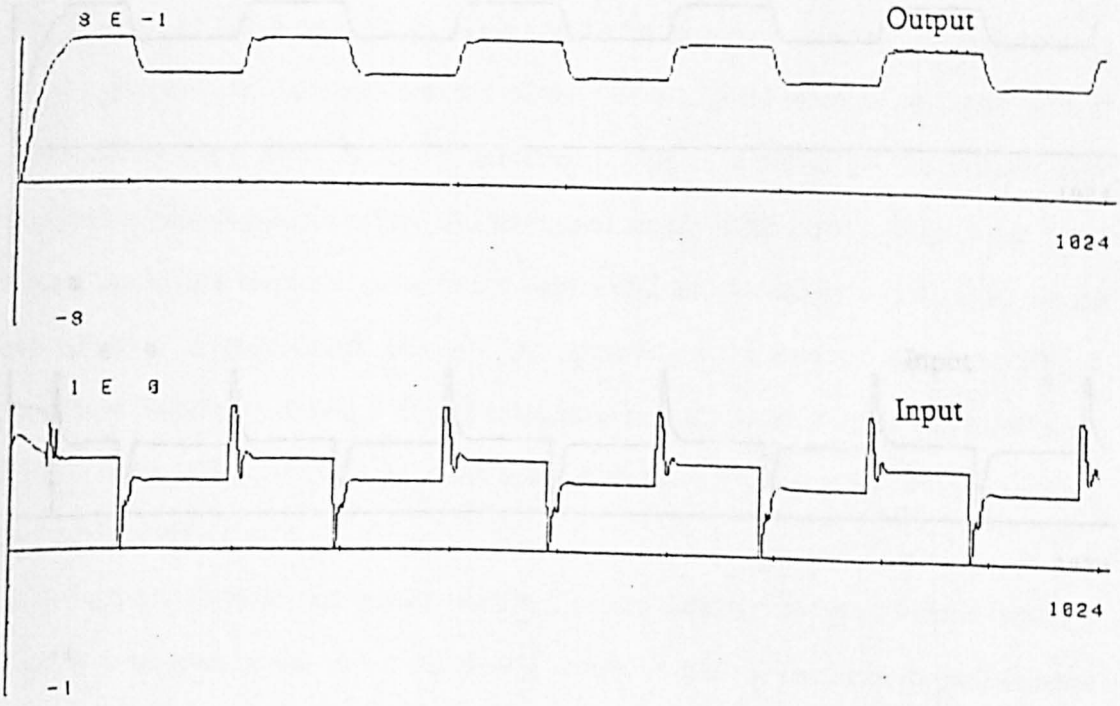


Figure 6.25: Control of the Atracurium model (NLI) using GPC

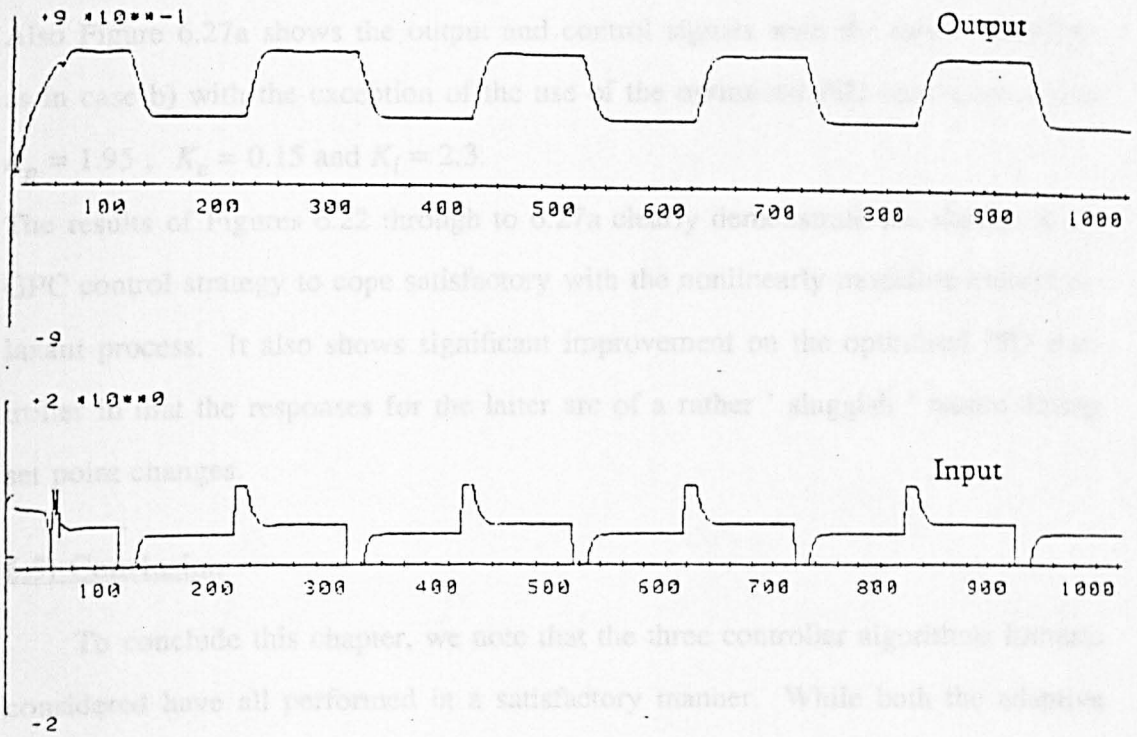


Figure 6.26: Control of the Atracurium model (NOI) using GPC

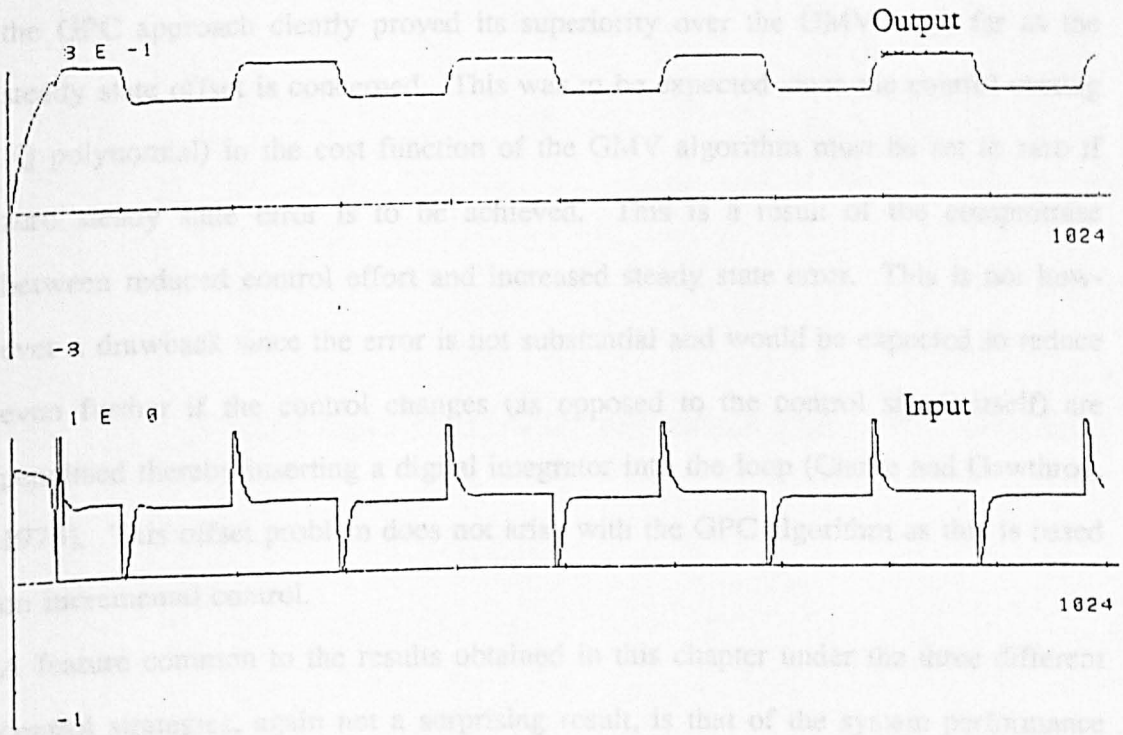


Figure 6.27: Control of the Atracurium model (NOI) using GPC

Also Figure 6.27a shows the output and control signals with the same conditions as in case b) with the exception of the use of the optimized PID controller, where $K_p = 1.95$, $K_e = 0.15$ and $K_i = 2.3$.

The results of Figures 6.22 through to 6.27a clearly demonstrate the ability of the GPC control strategy to cope satisfactory with the nonlinearly modelled muscle relaxant process. It also shows significant improvement on the optimised PID controller in that the responses for the latter are of a rather ' sluggish ' nature during set point changes.

6.5) Conclusion

To conclude this chapter, we note that the three controller algorithms hitherto considered have all performed in a satisfactory manner. While both the adaptive control schemes coped well with the system nonlinearity and showed faster responses during set-point level changes over the fixed parameter controller (PID), the GPC approach clearly proved its superiority over the GMV in so far as the steady state offset is concerned. This was to be expected since the control costing (Q polynomial) in the cost function of the GMV algorithm must be set to zero if zero steady state error is to be achieved. This is a result of the compromise between reduced control effort and increased steady state error. This is not however a drawback since the error is not substantial and would be expected to reduce even further if the control changes (as opposed to the control signal itself) are penalised thereby inserting a digital integrator into the loop (Clarke and Gawthrop, 1975). This offset problem does not arise with the GPC algorithm as this is based on incremental control.

A feature common to the results obtained in this chapter under the three different control strategies, again not a surprising result, is that of the system performance being of better quality the shorter the range of variation of the demand signal. As the requirement in muscle relaxation control is for a steady level of paralysis, a

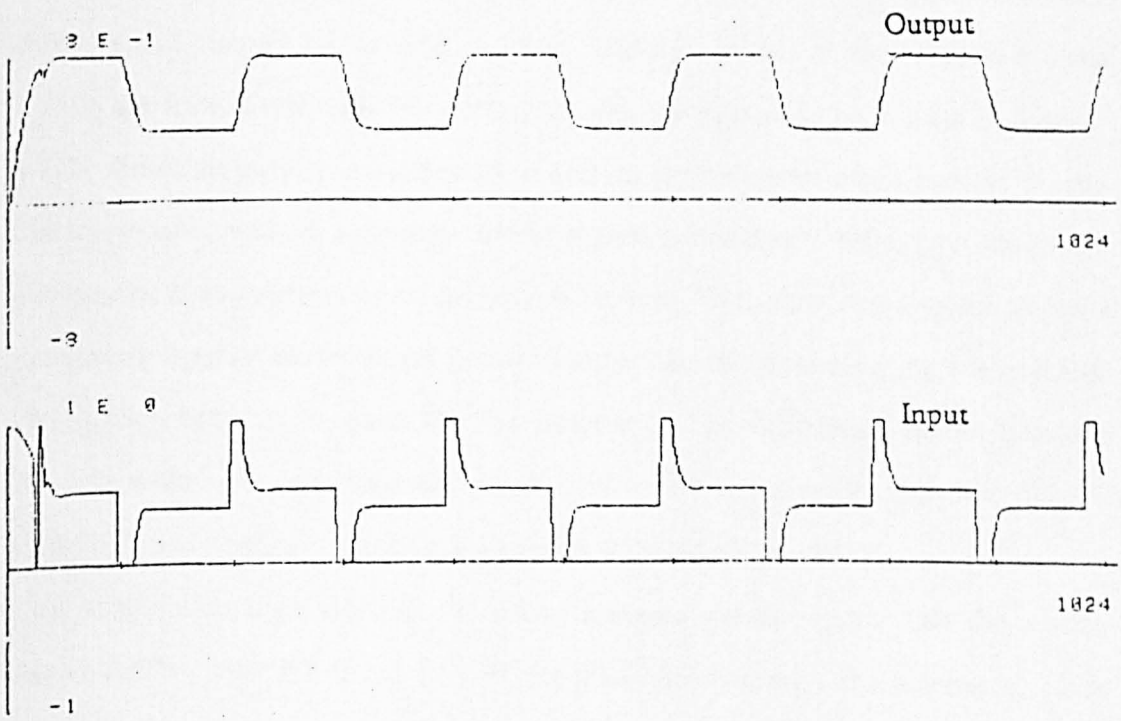


Figure 6.27a: Control of the Atracurium model (NOI) using GPC

range of set-point variation of 0.78-0.82, say for 80% degree of relaxation would be realistic and the above observation would encourage the GPC and GMV algorithms to be used in on-line practical situations.

An extention of the GMV self-tuner to non-linear systems modelled by nonlinear difference equations (Sales, 1988) is presented in the next chapter and applied to the muscle relaxant process.

CHAPTER 7

Nonlinear GMV Control Of Relaxant administration via simulation

7.1) Introduction

In the preceding chapter, the benefits of feedback control and its superiority to the conventional method (manual control) of muscle relaxant administration were highlighted. Reasons were given as to why preference was given to the adaptive (self-tuning) over the classical fixed parameter controller. The self-tuners described in chapter six as well as those used by Linkens et al (1985) and Denai (1988) are all derived with the assumption that the system to be control is linear. While these controllers are expected to and do perform well when applied to non-linear systems with a subtantial linear region, satisfactory behaviour cannot be guaranteed if the system's nonlinearity is severe. With care being taken to make the system operate in this linear portion (some kind of 'jacketting' regime (Menad, 1984) may need to be used for this purpose). The self-tuners of the previous chapter were seen to maintain good control of the muscle relaxant process modelled as a nonlinear difference equation (NARMAX model).

This observation together with the above remarks would suggest that the system was operating over the linear part of the pharmacodynamics characteristics. It is also true to say that the NARMAX models were not severely nonlinear in that recording of the data, from which they were identified, did not begin until the margin of safety (dead-space) was taken up.

In the light of the above remarks regarding the possible failure of linear self-tuners when applied to some nonlinear systems, there is certainly a case for having a self-tuner controller algorithm in which the on-line identification is performed on the basis of a nonlinear model of the process, and the subsequent control law computed using the knowledge thus required about the system's nonlinear dynamics.

The usefulness of nonlinear identification packages such as the NLI or NOI is evident as these provide an off-line means of determining the nonlinear structure (NARMAX model) of the system under consideration. Once a NARMAX model structure has thus been identified, it may be used in a self-tuning context with its parameters being estimated on-line and the desired control law computed.

In what follows, two adaptive control approaches of the kind described above are considered. The first is a direct digital control based on a k-step ahead prediction (k being the process time delay) and the second an extention of the GMV self-tuner of the previous chapter to nonlinear systems (Sales, 1988). In both cases, the system's nonlinearity is in form of a NARMAX model.

7.2) K-step ahead predictor control

Given a nonlinear difference equation relating the system's output z(t) to its past values and to those of the input (control) signal, and given that the system's integer input delay is k, the control signal u(t) is calculated by predicting z(t+1),.....,z(t+k-1) using the given difference equation shifted forward in time and setting z(t+k) equal to the desired output (i.e the set-point).

7.2.1) The basic algorithm

consider the nonlinear difference equation:

$$z(t) = F^l[z(t-1),....,z(t-n_z);u(t-k),....u(t-k-n_u+1);e(t-1),....,e(t-n_e)] \qquad (7.1)$$

On-line identification will yield the following model (estimated).

$$z(t) = \hat{F}^l[z(t-1),....,z(t-n_z);u(t-k),....u(t-k-n_u+1)] \qquad (7.2)$$

$\hat{F}^l$ being the nonlinear (estimated) input/output map. Shifting equation (7.2) k-steps ahead gives:

$$z(t+k) = \hat{F}^l[z(t+k-1),....,z(t+k-n_z);u(t),....u(t-n_u+1)] \qquad (7.3)$$

The k-step ahead controller, as mentioned above, consists in solving for u(t), the control signal, by setting z(t+k)=w(t), w(t) being the desired output.

Hence setting the left-hand side of equation 7.3 to w(t) and solving for u(t) yields;

$$u(t) = G[w(t);z(t+k-1),....,z(t+k-n_z);u(t-1),....u(t-n_u+1)] \tag{7.4}$$

where

G[.] is a nonlinear function of w, z and u.

at time t, z(t+1),.....,z(t+k-1) are all unknown and have to be predicted thus:

$$\left\{\begin{array}{l} \hat{z}(t+1) = \hat{F}^l[z(t),....,z(t+1-n_z);u(t-k+1),....,u(t-n_u-k+2)] \\ \hat{z}(t+2) = \hat{F}^l[z(t+1),....,z(t+2-n_z);u(t-k+2),....,u(t-n_u-k+3)] \\ \qquad\qquad\qquad\qquad \vdots \\ \hat{z}(t+k-1) = \hat{F}^l[\hat{z}(t+k-2),....,z(t+k-1-n_z);u(t-1),....,u(t-n_u)] \end{array}\right\} \tag{7.5}$$

The nonlinear k-step ahead adaptive control law is given by substitution of equation 7.5 into 7.4 thus:

$$u(t) = G[w(t);\hat{z}(t+k-1),....,z(t+k-n_z);u(t-1),....,u(t-n_u+1)] \tag{7.6}$$

7.2.2) K-step ahead predictor control of muscle relaxation

a) Case of the Vecuronium model

The NARMAX model used to simulate the process nonlinear dynamics is that of equation 4.5. The time delay in this case is 3 (i.e k=3) and so two values of z are predicted and the control law computed on the basis that z(t+3)=w(t), the set-point w(t) being a square wave. Figures 7.1 and 7.2 show the system responses.

In Figure 7.1 where no limits are imposed on the control signal, u(t) is seen to go both negative and 'excessively' positive during set point level changes. Both these features are unrealistic since negative control signals are equivalent to a sucking action i.e that of extracting the drug from the blood stream, while excessively positive controls may damage the pump motor if its input voltage upper limit is exceeded i.e for a 10 volt limit, the control signal should not exceed 1 when all the variables are normalised, which is the case for our application. The set-point tracking is nevertheless very good in terms of both transient and steady

state behaviours. Figures 7.2 shows the same variables as those in Figure 7.1 but with control limits of 0 and 1 being now imposed. At first sight, the control signal appears to be 'badly behaved' but as this happens between 'safe' limits with good steady state levels being kept everywhere except at the transient points, it can be concluded that the control is 'good'. One would expect it to be even more so in real-life application where fluctuations in the set-point, if at all required, will be very small.

The output response is again well behaved except for the initial transient which is rather slow (cf. figure 7.1). The limiting actions on the control signal are reflected in the output response at the transition points, but the time constants with which the output changes levels are fast enough, demonstrating good control in the presence of the above constraint (ie.$0 \leq u(t) \leq 1$).

b) Case of the Atracurium NLI model

Equation 5.4 is used, here, to simulate the system dynamics. The integer delay for this case is 4 and so 3 future values of z have to be predicted and z(t+4) set to w(t), again a square wave set-point. Figures 7.3 and 7.4 show the results for this case with the control limits (absent in Figure7.3) being introduced in Figure 7.4. Similar comments to those for the above Vecuronium case apply here, with this case showing an even better 'clipped' control signal and a subsequently better behaviour of the output response with faster transients. This is clearly due to the fact that when the control signal was not 'clipped' (Figure 7.3) its excursions at the transient points are of smaller amplitude than those in Figure 7.1.

c) Case of the Atracurium NOI model

Figures 7.5 and 7.6 show the control, set-point and output signals for this case in which the NARMAX model of equation 5.7 simulates the system's non-linear dynamics. As for case b) the integer delay, here, is 4.

An attempt to run the simulation with no control limits resulted in numerical

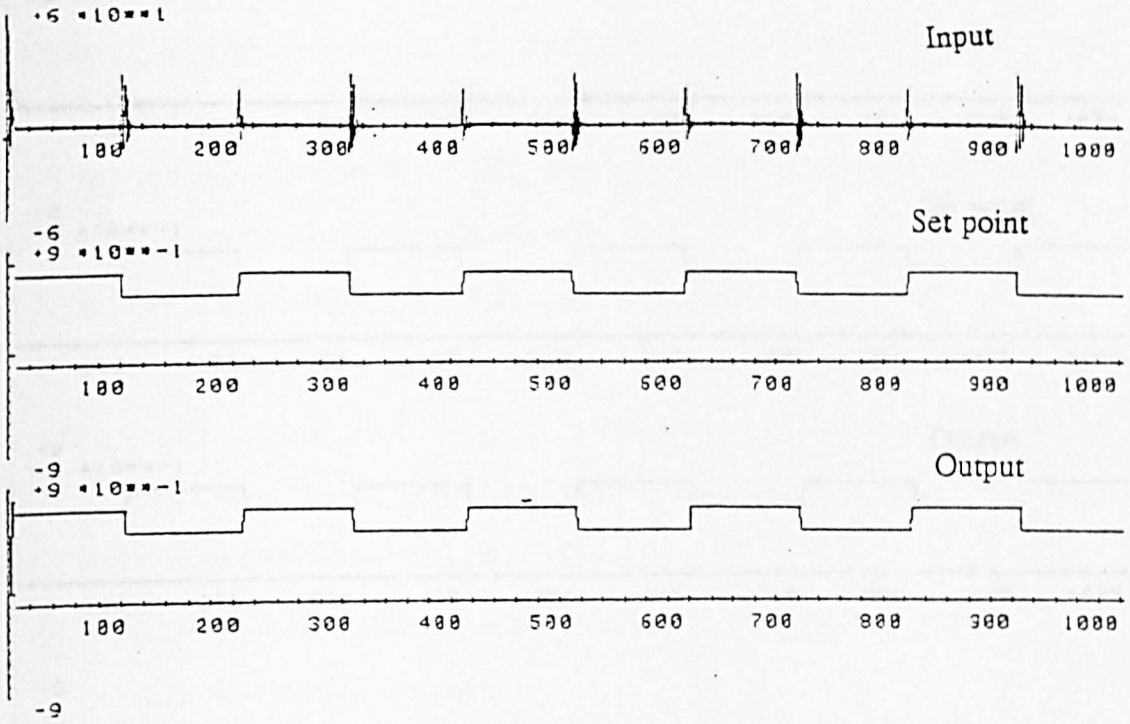


Figure 7.1: K-step ahead controlled system of Vecuronium

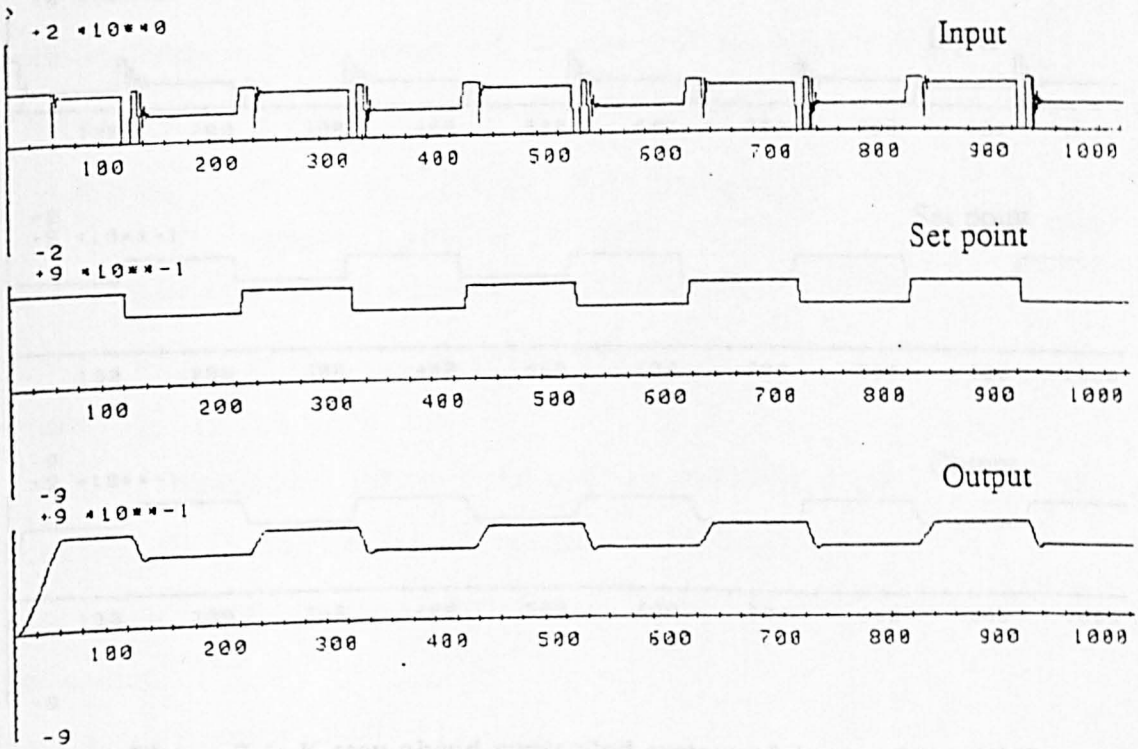


Figure 7.2: K-step ahead controlled system of Vecuronium

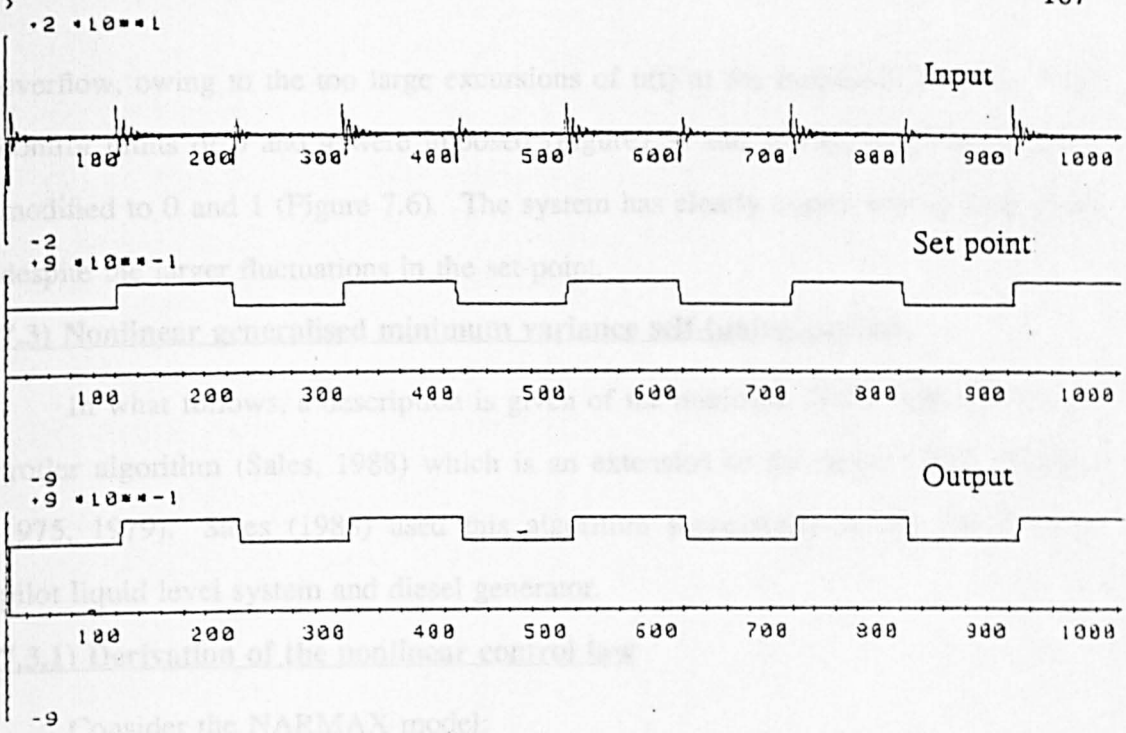


Figure 7.3: K-step ahead controlled system of Atracurium (NLI)

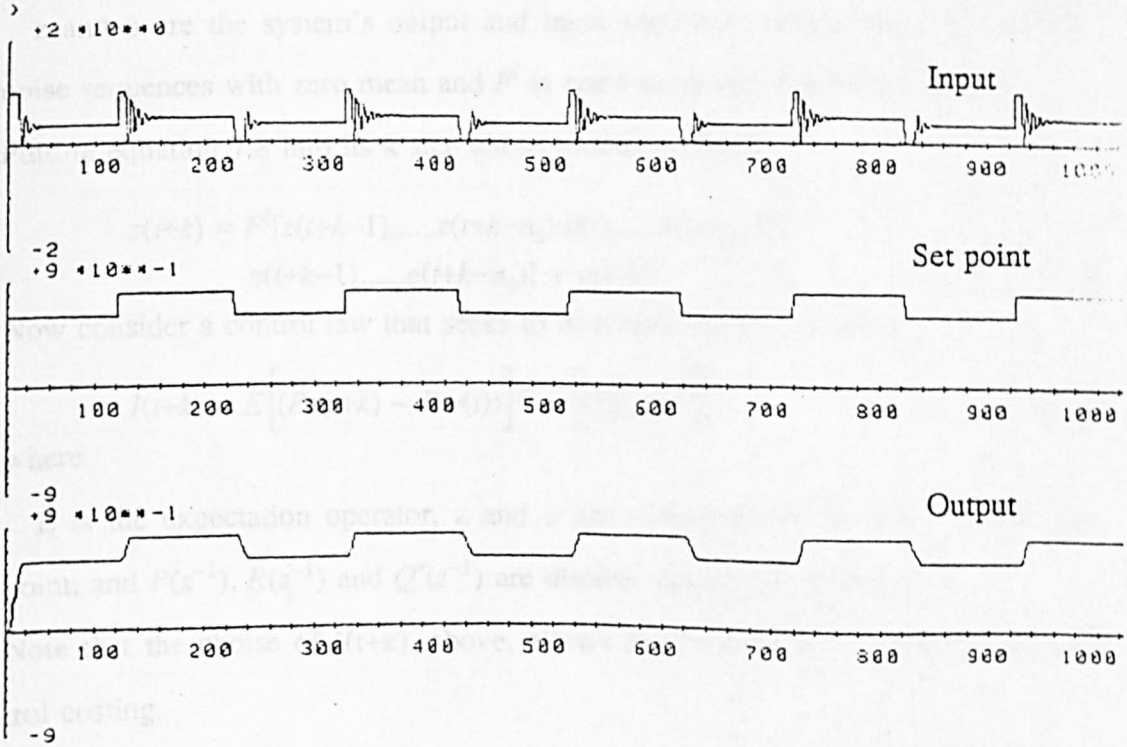


Figure 7.4: K-step ahead controlled system of Atracurium (NLI)

overflow, owing to the too large excursions of u(t) at the transition instants. First control limits of 0 and 4 were imposed (Figure7.5) and second, these limits were modified to 0 and 1 (Figure 7.6). The system has clearly coped well in both cases despite the larger fluctuations in the set-point.

7.3) Nonlinear generalised minimum variance self-tuning control

In what follows, a description is given of the nonlinear GMV self-tuning controller algorithm (Sales, 1988) which is an extension of the linear GMV (Clarke, 1975, 1979). Sales (1988) used this algorithm successfully in the control of a pilot liquid level system and diesel generator.

7.3.1) Derivation of the nonlinear control law

Consider the NARMAX model:

$$z(t) = F^l[z(t-1),....,z(t-n_z);u(t-k),....,u(t-k-n_u+1);\\ e(t-1),....,e(t-n_e)] + e(t) \tag{7.8}$$

where

z and u are the system's output and input sequences respectively, e(t) a white noise sequences with zero mean and F^l is some nonlinear function of degree l.

Putting equation 7.8 into its k-step ahead prediction form:

$$z(t+k) = F^l[z(t+k-1),....,z(t+k-n_z);u(t),....,u(t-n_u+1);\\ e(t+k-1),....,e(t+k-n_e)] + e(t+k) \tag{7.9}$$

Now consider a control law that seeks to minimise the cost function:

$$I(t+k) = E\Big[(Pz(t+k) - Rw(t))\Big]^2 + \Big[Q'f(u(t))\Big]^2 \tag{7.10}$$

where

E is the expectation operator, z and u are as previously defined, w is the set-point, and $P(z^{-1})$, $R(z^{-1})$ and $Q'(z^{-1})$ are discrete design polynomials.

Note that the choise of I(t+k), above, allows for both set-point tracking and control costing.

Substituting equation 7.9 into 7.10

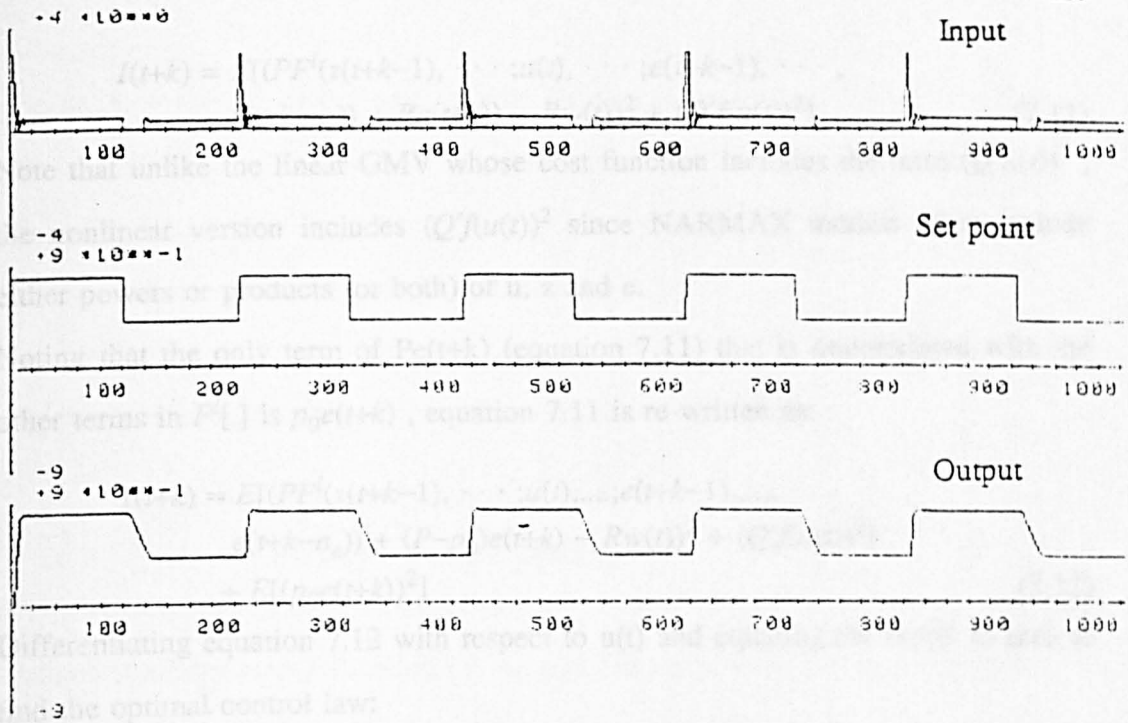


Figure 7.5: K-step ahead controlled system of Atracurium (NOI)

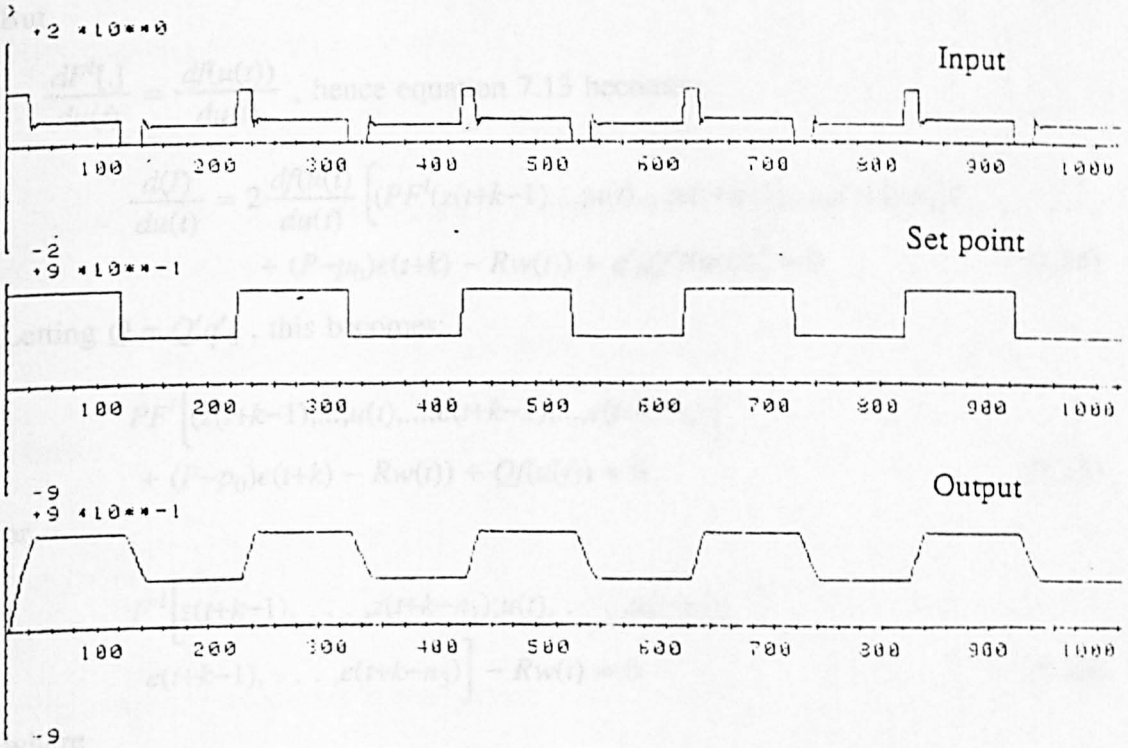


Figure 7.6: K-step ahead controlled system of Atracurium (NOI)

$$I(t+k) = E[(PF^l(z(t+k-1), \cdots ;u(t), \cdots ;e(t+k-1), \cdots ,$$
$$e(t+k-n_e)) + Pe(t+k)) - Rw(t))^2 + (Q'f(u(t))^2] \quad (7.11)$$

Note that unlike the linear GMV whose cost function includes the term $(Q'u(t))^2$, the nonlinear version includes $(Q'f(u(t))^2$ since NARMAX models often include either powers or products (or both) of u, z and e.

Noting that the only term of Pe(t+k) (equation 7.11) that is uncorrelated with the other terms in $F^l[.]$ is $p_0e(t+k)$, equation 7.11 is re-written as:

$$I(t+k) = E[(PF^l(z(t+k-1), \cdots ;u(t),....;e(t+k-1),....,$$
$$e(t+k-n_e)) + (P-p_0)e(t+k) - Rw(t))^2 + (Q'f(u(t))^2]$$
$$+ E[(p_0e(t+k))^2] \quad (7.12)$$

Differentiating equation 7.12 with respect to u(t) and equating the result to zero to find the optimal control law:

$$\frac{d(I)}{du(t)} = 2\frac{dF^l[.]}{du(t)}\Big[(PF^l(z(t+k-1),...;u(t),...;e(t+k-1),...,$$
$$e(t+k-n_e) + (P-p_0)e(t+k) - Rw(t))$$
$$+ 2q'_0Q'\frac{df(u(t))}{du(t)}f(u(t)) = 0 \quad (7.13)$$

But

$\frac{dF^l[.]}{du(t)} = \frac{df(u(t))}{du(t)}$, hence equation 7.13 becomes:

$$\frac{d(I)}{du(t)} = 2\frac{df(u(t)}{du(t)}\Big[(PF^l(z(t+k-1),...;u(t),...;e(t+k-1),...,e(t+k-n_e))$$
$$+ (P-p_0)e(t+k) - Rw(t)) + q'_0Q'f(u(t))\Big] = 0 \quad (7.14)$$

Letting $Q = Q'q'_0$, this becomes;

$$PF^l\Big[(z(t+k-1),...;u(t),...;e(t+k-1),...,e(t+k-n_e)\Big]$$
$$+ (P-p_0)e(t+k) - Rw(t)) + Qf(u(t)) = 0 \quad (7.15)$$

or

$$F'^l\Big[z(t+k-1), \ldots ,z(t+k-n_1);u(t), \ldots ,u(t-n_2);$$
$$e(t+k-1), \ldots ,e(t+k-n_3)\Big] - Rw(t) = 0 \quad (7.16)$$

where

$n_1 = n_z + n_p$

$n_2 = (n_u + n_q - 1)$ or $(n_u + n_p - 1)$ which ever is the greater

and

$n_3 = n_e + n_p$.

Next consider the auxiliary output function given by

$$\Phi(t+k) = Pz(t+k) - Rw(t) + Qf(u(t)) \tag{7.17}$$

Which, upon substitution for z(t+k) from equation 7.9, becomes

$$\Phi(t+k) = PF^{l}\Big[z(t+k-1),...;u(t),...;e(t+k-1), \ldots ,e(t+k-n_e)\Big] + Pe(t+k) - Rw(t) + Qf(u(t)) \tag{7.18}$$

It is now required to find a control law which minimises the variance of this generalised auxiliary output function i.e the cost function to be minimised is now

$$J(t+k) = E[(\Phi(t+k))^2] \tag{7.19}$$

Substituting equation 7.18 into 7.19 to give

$$\begin{aligned} J(t+k) = E[&(PF^{l}(z(t+k-1),...;u(t),...;e(t+k-1), \ldots ,e(t+k-n_e)) \\ &+ (P - p_0)e(t+k) - Rw(t) + Qf(u(t))^2] \\ &+ E[(p_0 e(t+k))^2] \end{aligned} \tag{7.20}$$

Differentiating with respect to u(t)

$$\frac{dJ}{du(t)} = 2(1+q_0)\frac{dF[.]}{du(t)}\Big[PF^{l}(z(t+k-1),...;u(t),...; e(t+k-1), \ldots ,e(t+k-n_0)) + (P-p_0)e(t+k) - Rw(t) + Qf(u(t))\Big] = 0 \tag{7.21}$$

and the control law bcomes

$$\begin{aligned} &F'^{l}[z(t+k-1), \ldots ,z(t+k-n_1),...;u(t), \ldots ,u(t-n_2); \\ &e(t+k-1), \ldots ,e(t+k-n_3)] - Rw(t) = 0 \end{aligned} \tag{7.22}$$

where

$n_1 = n_z + n_p$

$n_2 = (n_u + n_q - 1)$

or

$(n_u + n_p - 1)$ whichever is the greater

and

$n_3 = n_e + n_p$.

The control laws of equations 7.16 and 7.22 are identical, implying that if an auxiliary output is generated using equation 7.17, then it suffices to minimise its variance for a control law which minimises the more complex cost function (equation 7.10) to be arrived at.

So far in this section, all the derivations were on the basis of known system parameters, which is obviously not the case in self-tuning. So, let us now look at the nonlinear GMV algorithm, above, in the self-tuning context. There are two main cases to consider here i.e those of the algorithm being of an explicit or an implicit nature.

7.3.2) The explicit nonlinear GMV algorithm

The explicit self-tuning nonlinear GMV controller algorithm estimates the system parameter explicitly and uses them to compute the control law at every sampling instant.

The NARMAX model representation of the system (equation 7.8) is written in the form

$$z(t) = X^T(t)\Theta + e(t) \tag{7.23}$$

where the measurement vector $X^T(t)$ is given by

$$\begin{aligned} X^T(t) = [&z(t-1), \ldots, z(t-n_z); u(t-k), \ldots, u(t-k-n_u+1); \\ &e(t-1), \ldots, e(t-n_e); z^2(t-1), \ldots, z(t-1)u(t-k) \\ &, \ldots, z(t-1)e(t-1), \ldots \textit{ higher order terms}] \end{aligned}$$

and the parameter vector Θ by

$$\begin{aligned} \Theta = [&\Theta_1, \ldots, \Theta_{nz}; \Theta_{nz+1}, \ldots, \Theta_{nz+nu}; \Theta_{nz+nu+1}, \ldots, \Theta_{nz+nu+ne}; \\ &\Theta_{11}, \ldots, \Theta_{1nz+1}; \ldots\ldots\ldots, \Theta_{1nz+nu+ne} \ldots] \end{aligned}$$

Note that the noise parameters are estimated on-line i.e the algorithm should cope even with coloured noise corruption as long as the signal to noise ratio is high enough.

With the above definitions of X^T and Θ , the recursive least squares estimator of chapter 3 (equations 3.13 to 3.17) may be used to estimate the vector of parame-

ters Θ . The prediction error is now: [$z(t)-x^T(t)\ \hat{\Theta}(t-1)$], Θ , being the vector of parameter estimates. The noise terms in $X^T(t)$ cannot be measured and are hence estimated thus:

$$\hat{e}(t) = z(t) - \hat{X}(t)\hat{\Theta}(t).$$

7.3.3) The implicit nonlinear GMV self-tuner

Unlike the explicit algorithm, this estimates directly the controller parameters. The system parameters are thus indirectly or (implicitly) estimated.

For this case consider the auxiliary output function

$$\Phi(t) = Pz(t) - Rw(t-k) + Qf(u(t-k)) \tag{7.24}$$

$$\Phi(t) = F'^{l}[z(t-1),..,z(t-n_1);u(t-k), \ldots ,u(t-k-n_2);e(t-1), \ldots ,e(t-n_3)] - Rw(t-k) + e(t) \tag{7.25}$$

Re-write this as

$$\Phi(t) = X^T(t)\Theta + e(t) \tag{7.26}$$

with

$$X^T(t) = [z(t-1), \ldots ,z(t-n_1);u(t-k), \ldots ,u(t-k-n_2);e(t-1), \ldots , e(t-n_3);w(t-k),....,w(t-k-n_r);z(t-1)^2,...;z(t-1)z(t-2), \ldots , z(t-1)u(t-k),...;z(t-1)e(t-1),...;.... \textit{higher order terms}] \tag{7.27}$$

and

$$\Theta = [\Theta_1, \ldots ,\Theta_{n1};\Theta_{n1+1},....\Theta_{n1+n2},..;\Theta_{n1+n2+1}, \ldots ,\Theta_{n1+n2+n3}; r_0,...r_{nr};\Theta_{11},....\Theta_{1n1};....\Theta_{1n1+n2+n3},....;....]$$

As for the explicit case, the recursive least squares estimator (equations 3.13-3.17) can be used, on-line, to estimate directly the controller parameters. The prediction error for this case is; ($\varepsilon(t) = \Phi(t) - X^T\Theta(t-1)$) and the noise term estimates are given by

$$\hat{e}(t) = \Phi(t) - X^T(t)\Theta(t)$$

The size of the measurement and parameters vectors is clearly larger in the case of the implicit algorithm, and for this reason, the explicit approach is adopted in the sequel.

7.4) Nonlinear GMV self-tuning control of muscle relaxation

a) Case of the Vecuronium

a.1) the simulation is run over 1000 samples with the following parametrisation of the self-tuning algorithm

$P(z^{-1}) = R(z^{-1}) = 1$; $Q(z^{-1}) = \lambda = 0.5$, k=3, the initial diagonal elements of the covariance matrix are set to 10^8 as this case is noise-free. The set-point is a square wave of levels 0.4 and 0.7 and the control signal is limited to between 0 and 1.

Figure 7.7 shows the control, set-point and output signals, while the estimated system parameters are displayed in Figures 7.8 through to 7.11. Note that only the first 200 samples are shown in the graphs of the parameters estimates, the reason being that these converge to constant values even before the 200th sample.

Figure 7.7 demonstrates a well-behaved system with good transient and steady state properties. The parameter estimates take between 10 to 20 samples to converge to their correct values, except for the last two which take about 150 samples to converge. This has not, however, influenced the control much. Note that the initial vector of parameter estimates Θ is taken to be null vector. This would obviously pose a problem in the computation of the control law initially as this includes division by a term that will be zero. To avoid such an occurrence, a statement is included in the control program to set this term to an arbitrary value of 0.1 when it is zero. This precaution can be done away with if a fixed-term controller is used to assume initial control and the estimation is allowed to run from t=0.

a.2) So next, a.1) is re-run with the initial control being provided by a PI controller, until the output reaches 75% of the first set-point level. The PI parameters are set to:

K_p=1.0 and K_i=0.08.

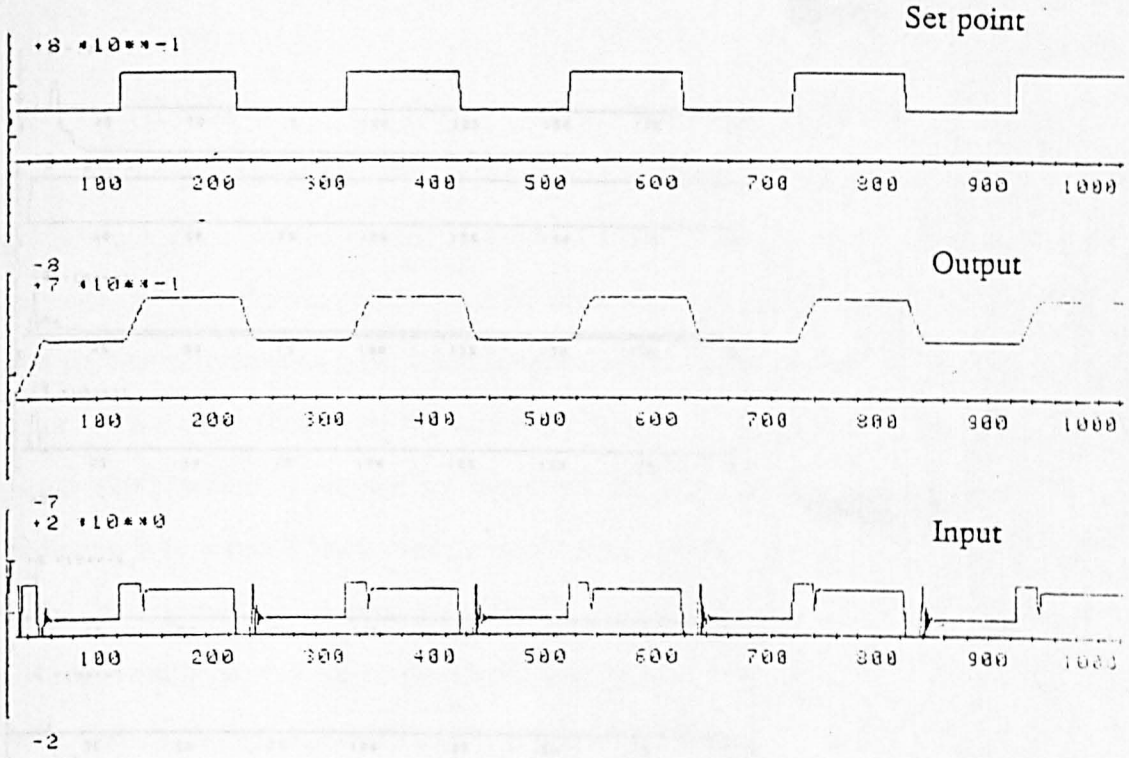


Figure 7.7: Nonlinear GMV controlled system of Vecuronium

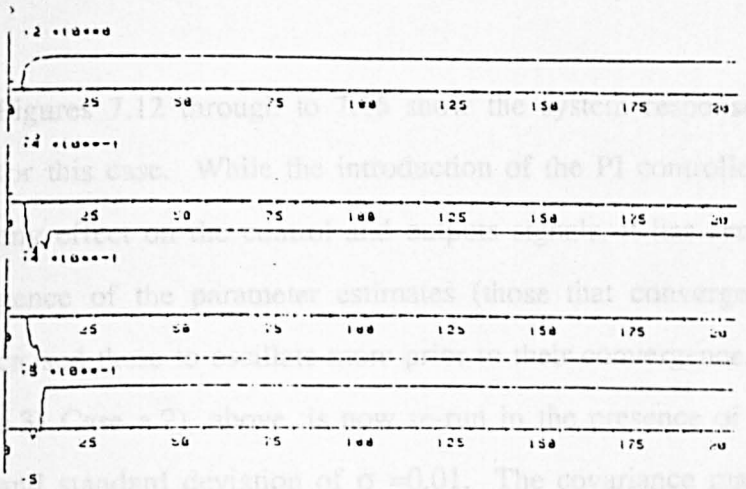


Figure 7.8

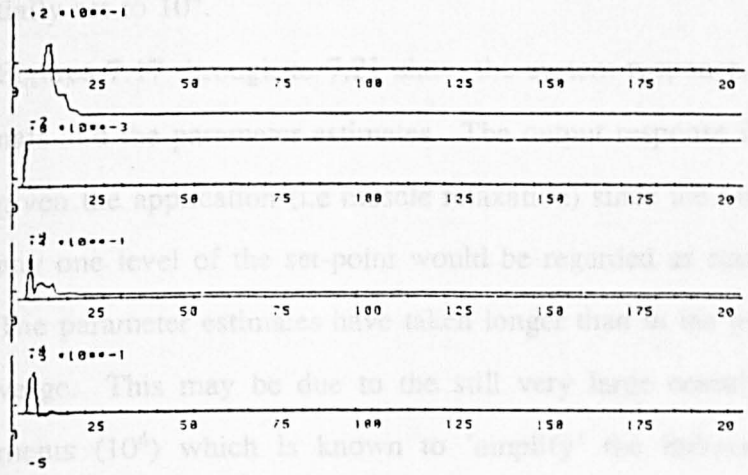


Figure 7.9

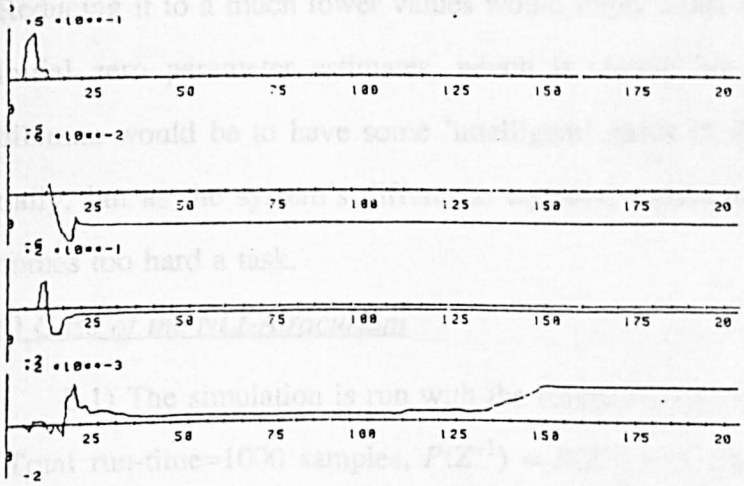


Figure 7.10

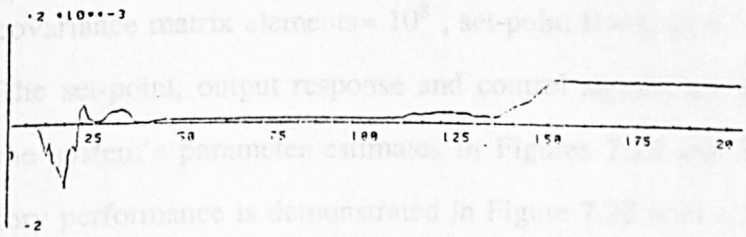


Figure 7.11

Figures 7.8-7.11: Parameter estimates

Figures 7.12 through to 7.16 show the system response and parameter estimates for this case. While the introduction of the PI controller does not appear to have any effect on the control and outputs signals, it has certainly delayed the convergence of the parameter estimates (those that converged fast in case a.1)) and caused these to oscillate more prior to their convergence.

a.3) Case a.2), above, is now re-run in the presence of white noise of zero mean and standard deviation of σ =0.01. The covariance matrix elements are now initially set to 10^4.

Figures 7.17 through to 7.21 show the system response, control and set-point signals and the parameter estimates. The output response is reasonably well behaved given the application (i.e muscle relaxation) since the observed fluctuations around any one level of the set-point would be regarded as acceptable by most surgeons. The parameter estimates have taken longer than in the preceding two cases to converge. This may be due to the still very large covariance matrix diagonal elements (10^4) which is known to 'amplify' the influence of noisy disturbances. Reducing it to a much lower values would imply some degree of confidence in the initial zero parameter estimates, which is clearly not true. A way out of this dilemna would be to have some 'intelligent' guess of the parameter estimates initially, but as the system's difference equation representation is nonlinear this becomes too hard a task.

b) Case of the NLI-Atracurium

b.1) The simulation is run with the following specification:

Total run-time=1000 samples, $P(Z^{-1})$ = $R(Z^{-})$ = 1, $Q(Z^{-1})=\lambda$ = 0.5 , k=4, initial covariance matrix elements= 10^8 , set-point levels (0.4, 0.7), control limits=(0, 1). The set-point, output response and control signals are shown in Figures 7.22 and the system's parameter estimates in Figures 7.23 and 7.24. Again very satisfactory performance is demonstrated in Figure 7.22 with a less oscillatory control

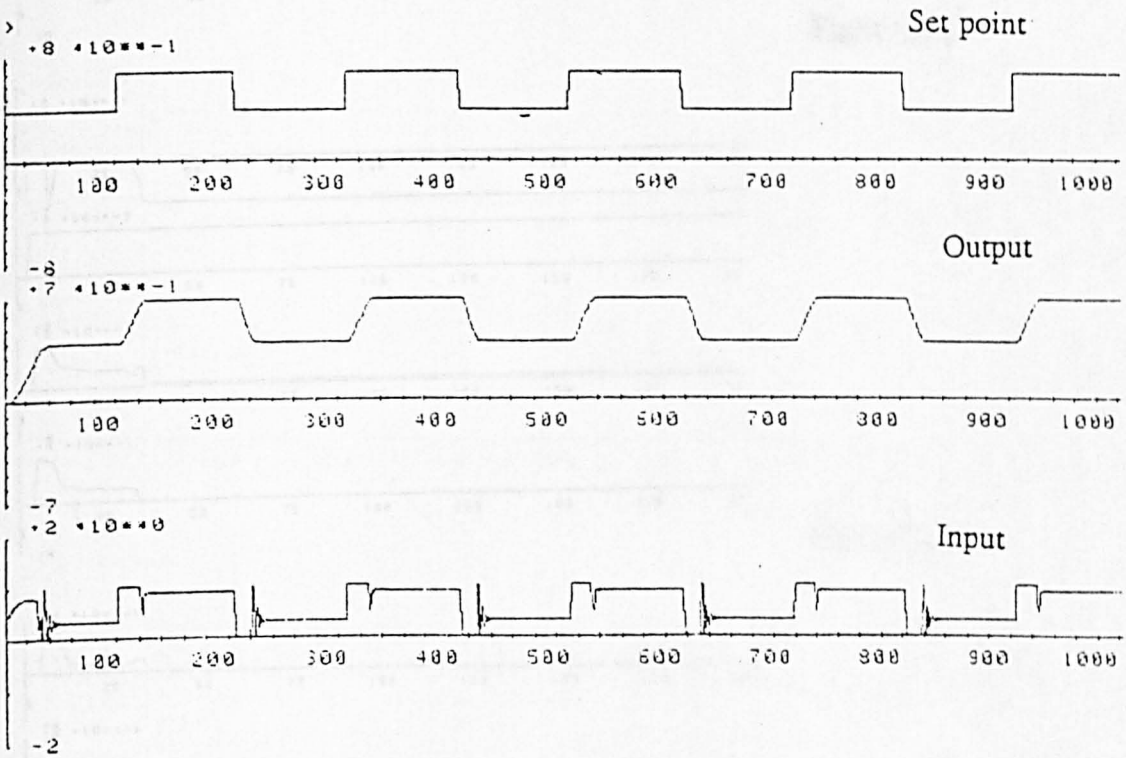


Figure 7.12: Nonlinear GMV controlled system of Vecuronium with a PI

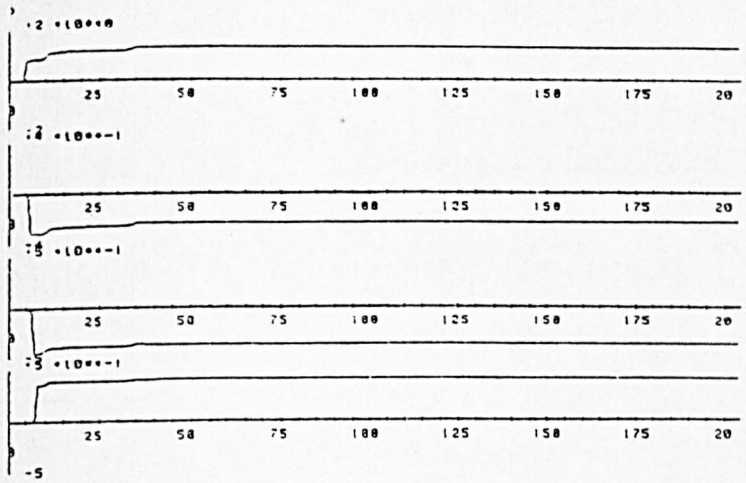


Figure 7.13

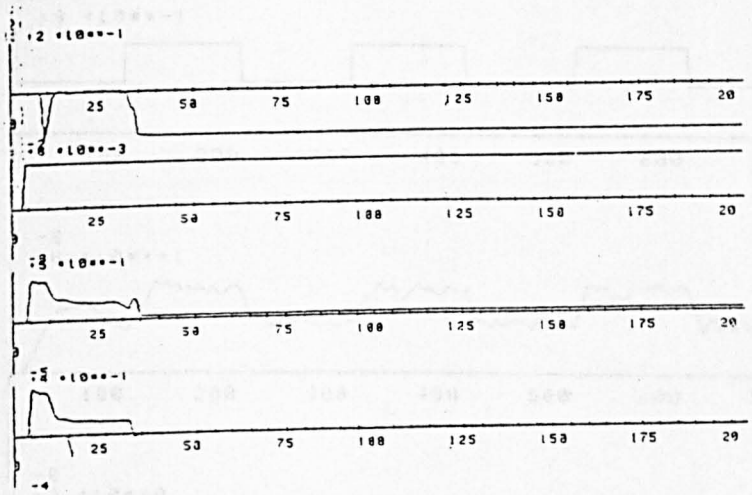


Figure 7.14

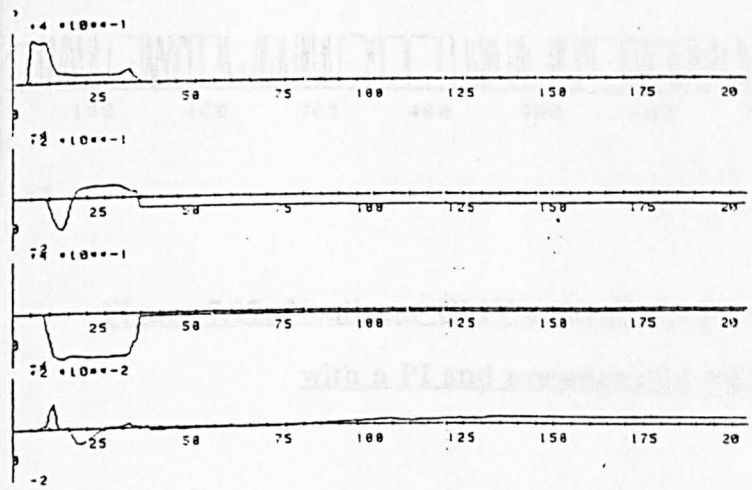


Figure 7.15

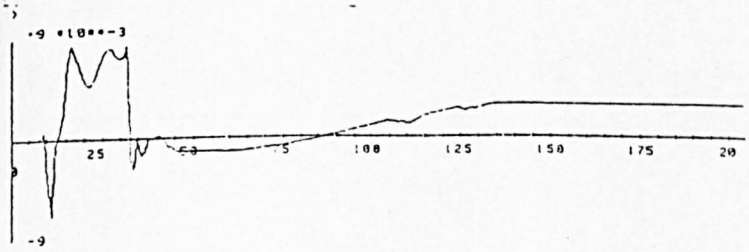


Figure 7.16

Figures 7.13-7.16: Parameter estimates

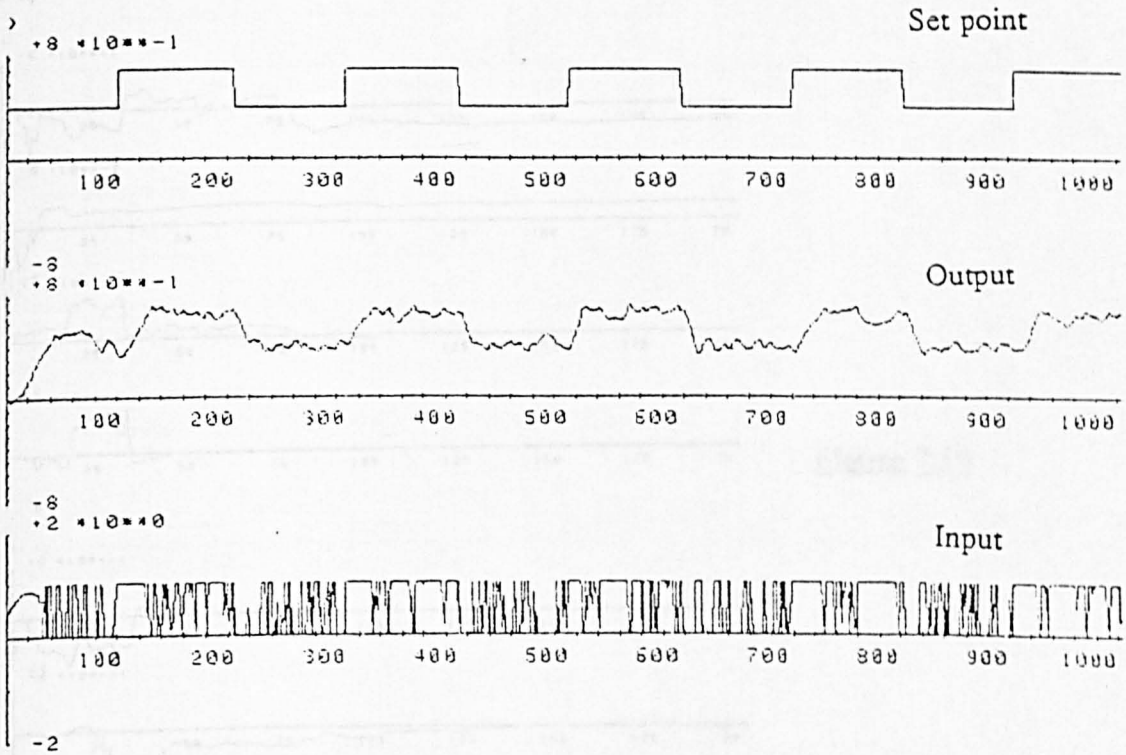


Figure 7.17: Nonlinear GMV controlled system of Vecuronium with a PI and presence of a white noise

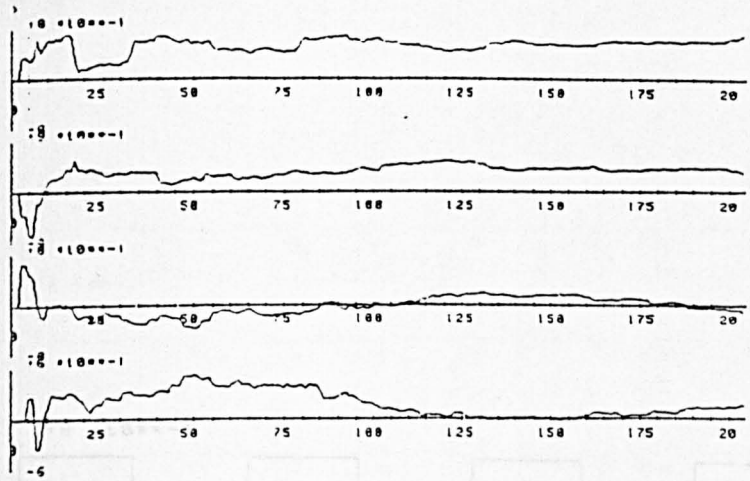


Figure 7.18

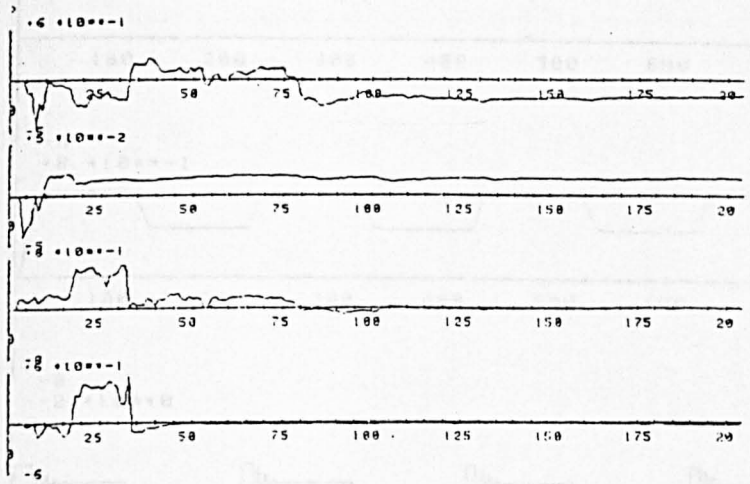


Figure 7.19

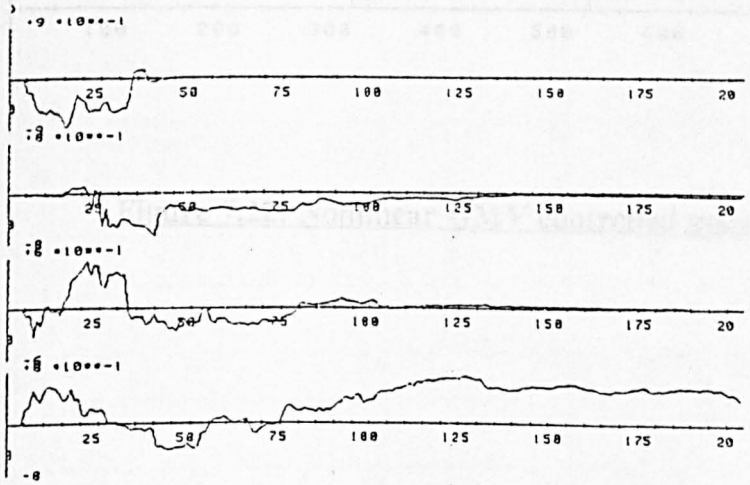


Figure 7.20

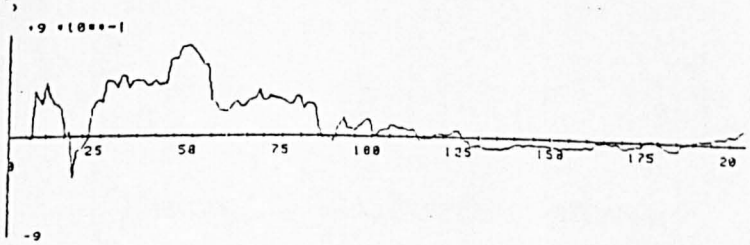


Figure 7.21

Figures 7.18-7.21: Parameter estimates

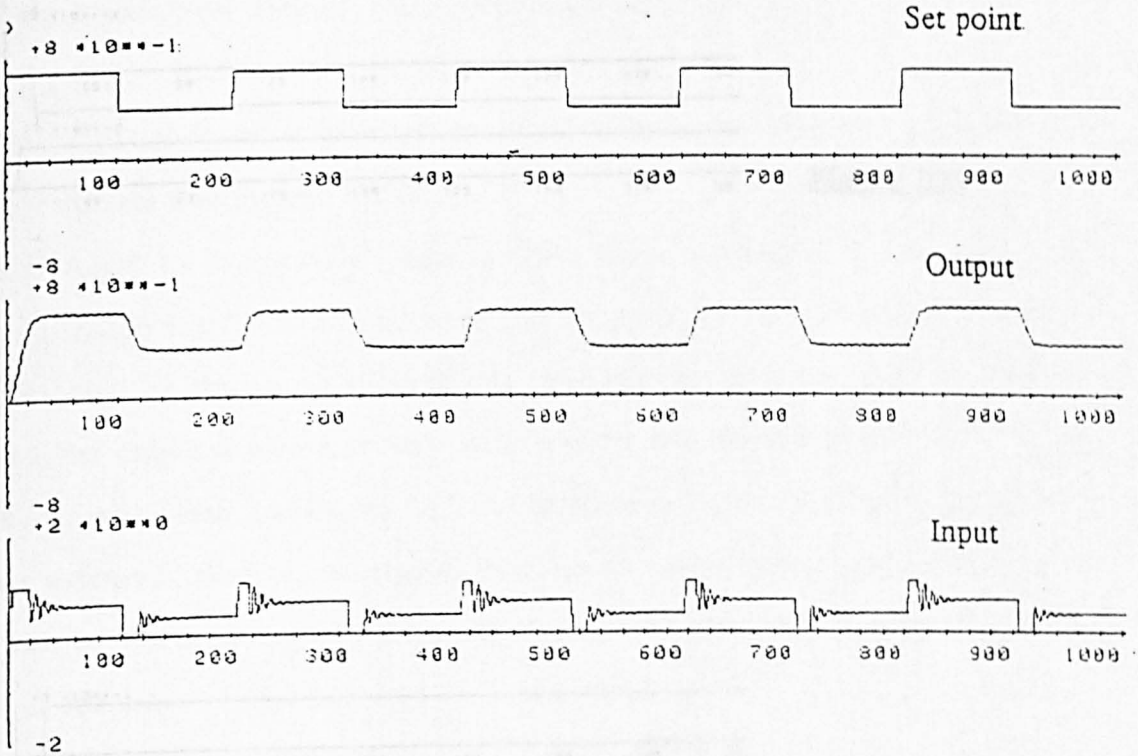


Figure 7.22: Nonlinear GMV controlled system of Atracurium (NLI)

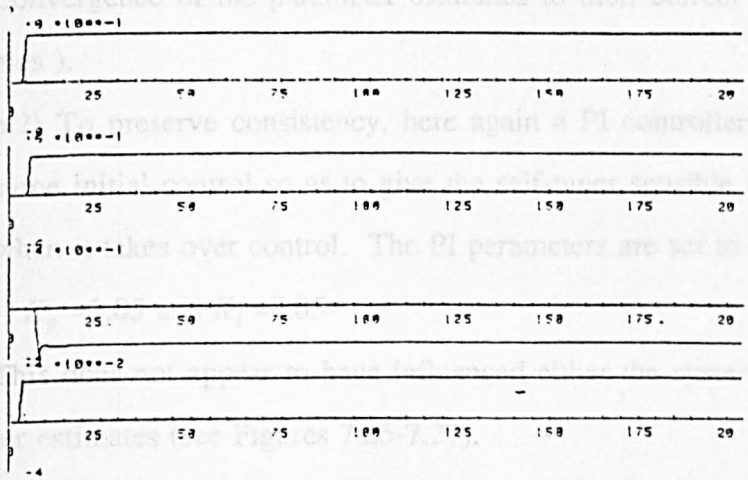


Figure 7.23

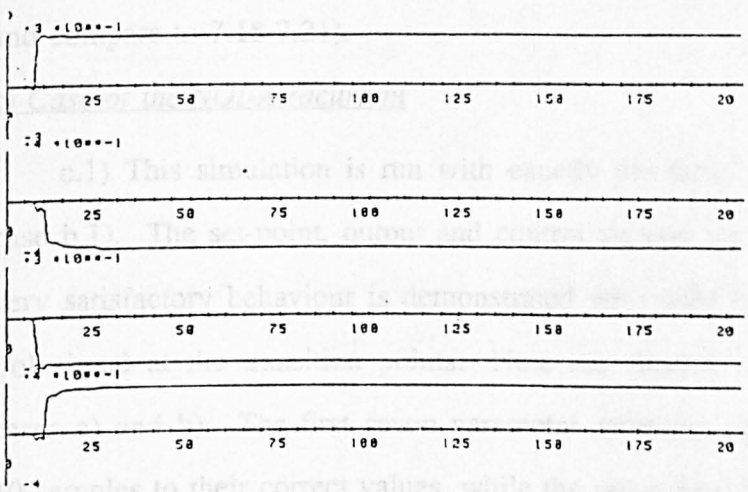


Figure 7.24

Figures 7.23-7.24: Parameter estimates

signal at the transient points (cf. Figure 7.7) than that in the Vecuronium case. Convergence of the parameter estimates to their correct values is fast (≤10 samples).

b.2) To preserve consistency, here again a PI controller is now introduced to assume initial control so as to give the self-tuner sensible initial parameter estimates when it takes over control. The PI parameters are set to:

K_p =1.05 and K_i =0.05

This does not appear to have influenced either the system response or the parameter estimates (see Figures 7.25-7.27).

b.3) Again for consistency's sake, a white noise disturbance of zero mean and 0.01 standard deviation is introduced and the covariance matrix diagonal elements set to 10^4. Although the effect of this noise is quite apparent on the control signal, the output response is very well-behaved and certainly better than in the Vecuronium case (see Figure 7.28 and compare to 7.17). Here again the parameter estimates failed to converge to their correct values though their initial oscillations are not as severe as those in the Vecuronium case (see Figures 7.29 and 7.30 and compare to 7.18-7.21).

c) Case of the NOI-Atracurium

c.1) This simulation is run with exactly the same specifications as those for case b.1). The set-point, output and control signals are shown in Figure 7.31. A very satisfactory behaviour is demonstrated with only sudden changes in the control signal at the transition points. Note the absence of any oscillations, unlike cases a) and b). The first seven parameter estimates have converged after about 10 samples to their correct values, while the remaining five have taken around 200 samples to converge (Figures 7.32 through 7.34).

c.2) Introduction of the PI controller with

K_p =1.05 and K_i =0.05

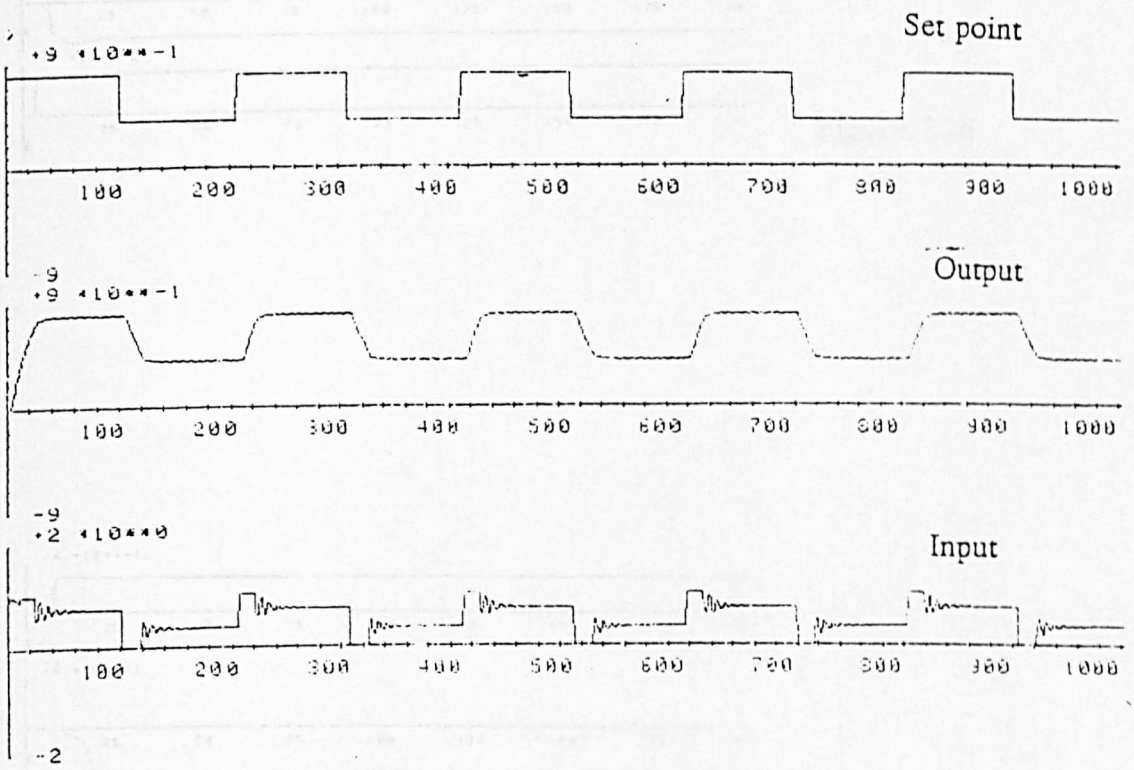


Figure 7.25: Nonlinear GMV controlled system of Atracurium (NLI) with a PI

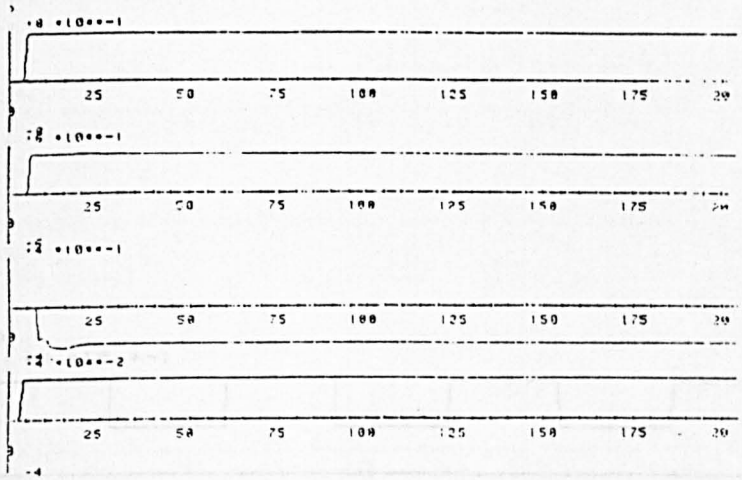


Figure 7.26

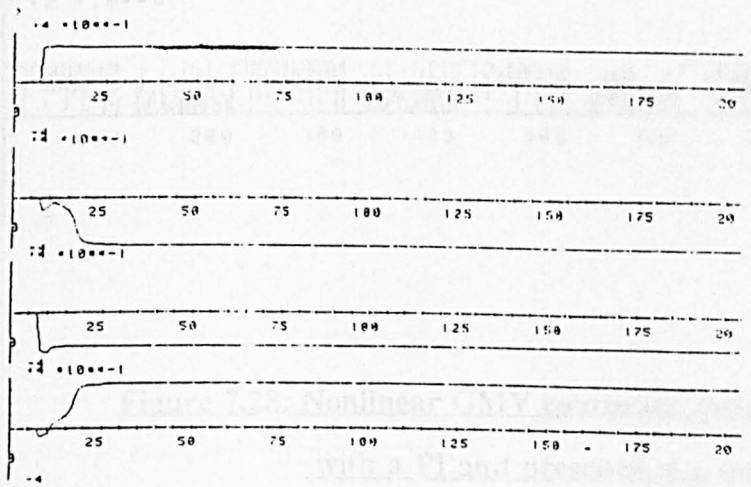


Figure 7.27

Figures 7.26-7.27: Parameter estimates

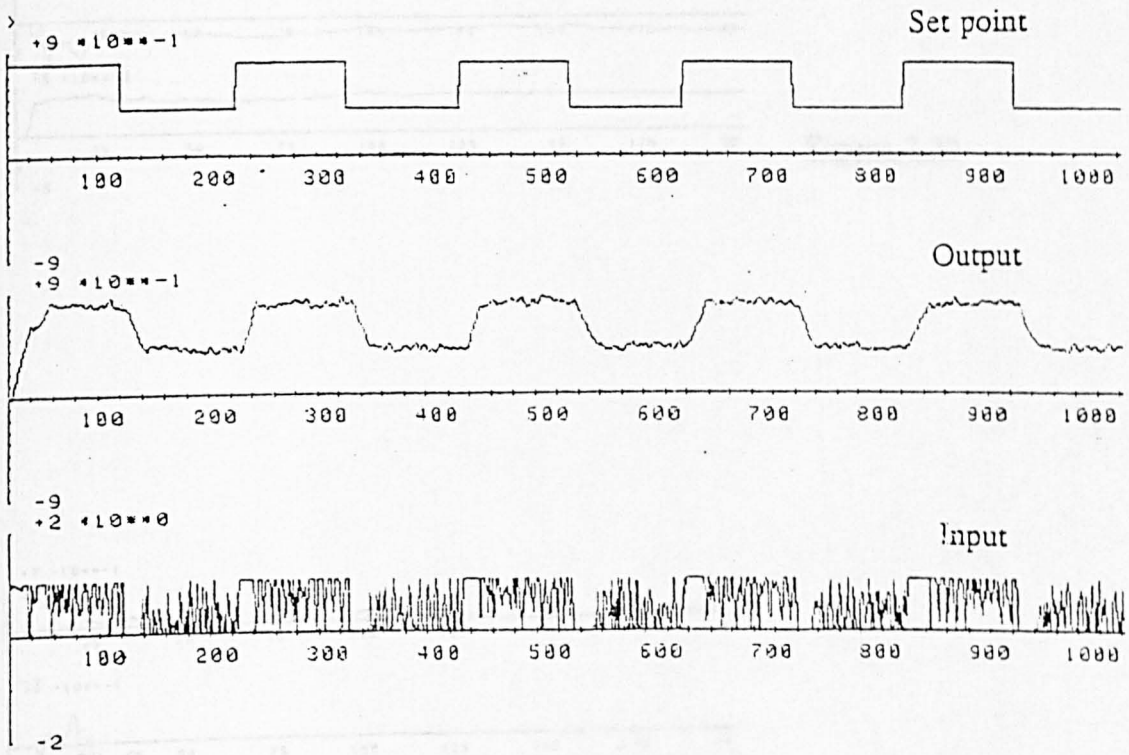


Figure 7.28: Nonlinear GMV controlled system of Atracurium (NLI) with a PI and presence of a white noise

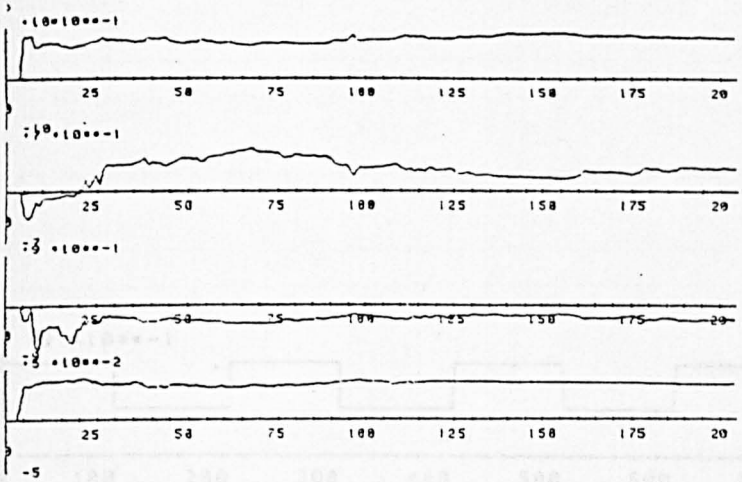


Figure 7.29

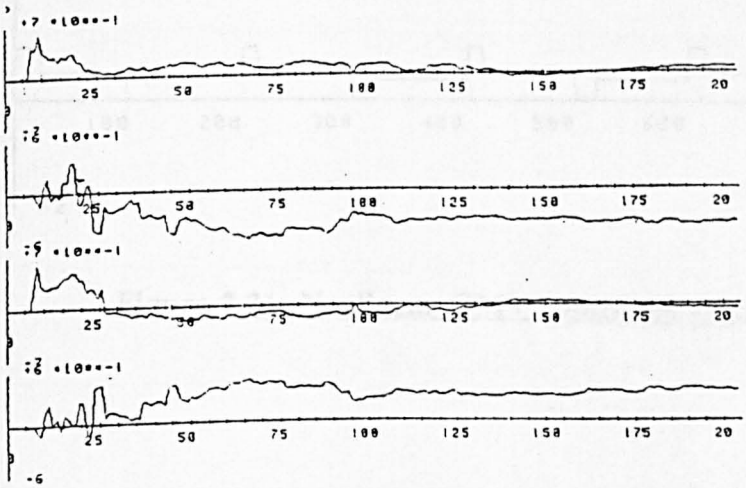


Figure 7.30

Figures 7.29-7.30: Parameter estimates

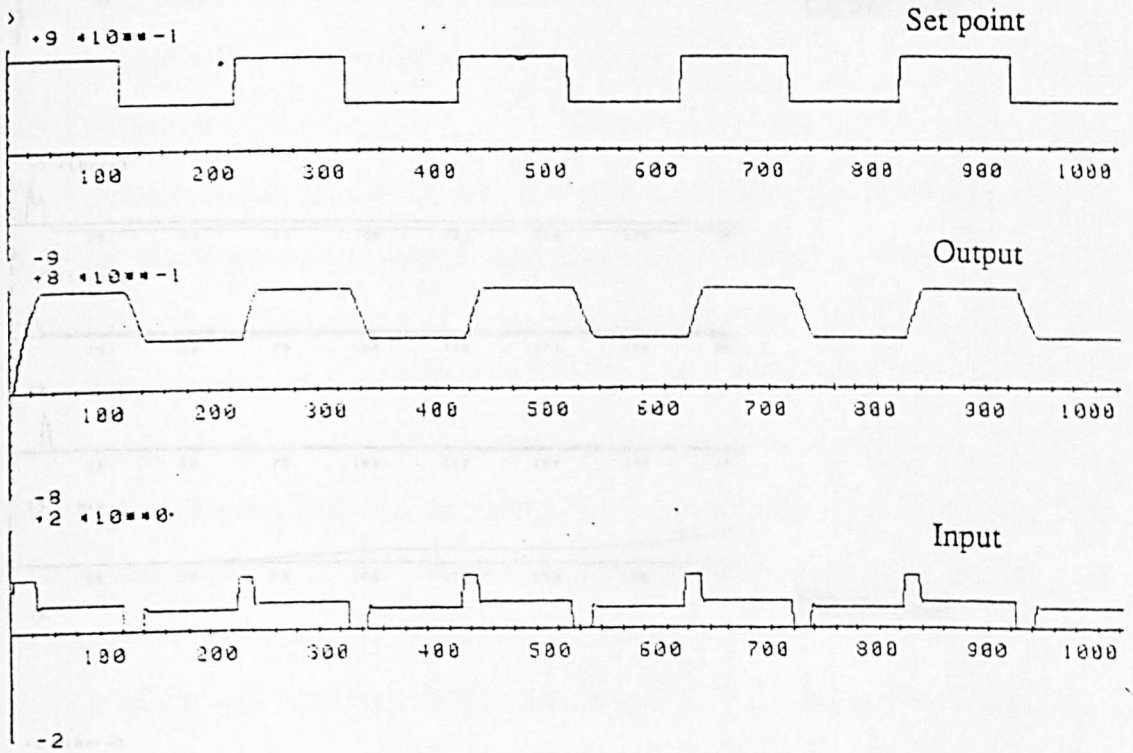


Figure 7.31: Nonlinear GMV controlled system of Atracurium (NOI)

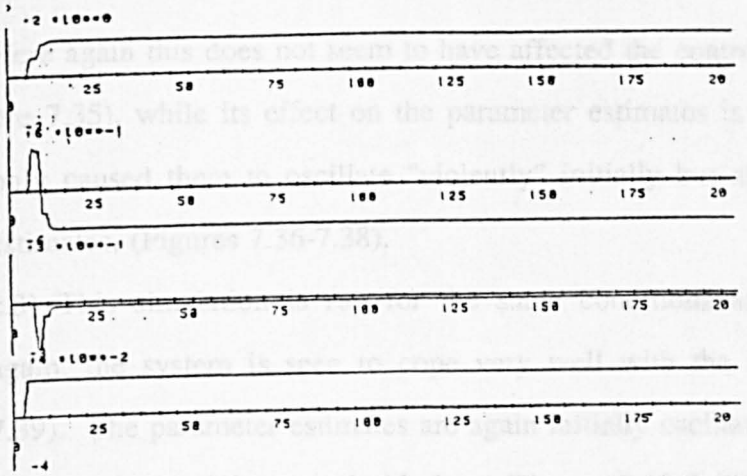


Figure 7.32

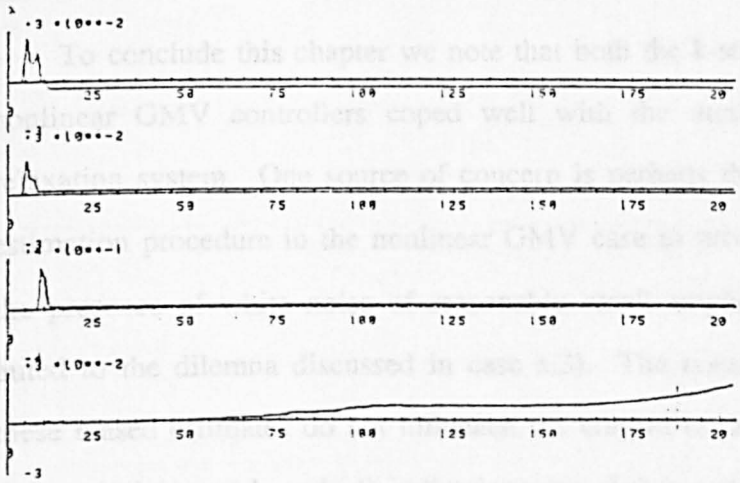


Figure 7.33

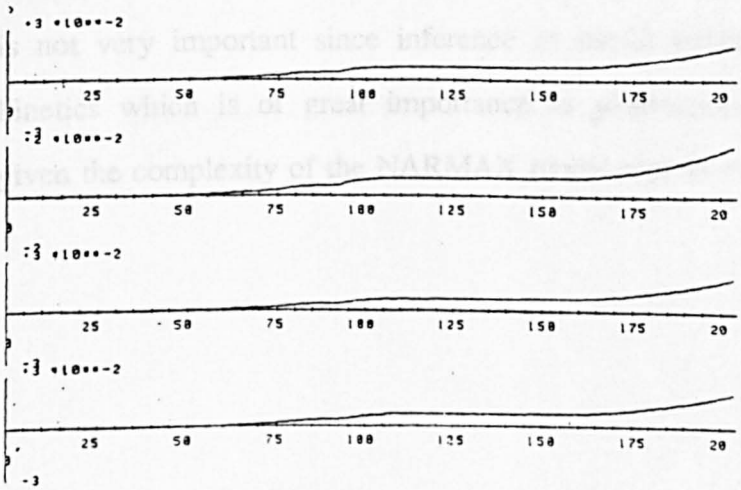


Figure 7.34

Figures 7.32-7.34: Parameter estimates

Here again this does not seem to have affected the control and output signals (Figure 7.35), while its effect on the parameter estimates is totally unexpected: it not only caused them to oscillate "violently" initially but also to converge to biased estimates, (Figures 7.36-7.38).

c.3) This simulation is run for the same conditions as case b.3) above. Here again, the system is seen to cope very well with the noisy disturbance (Figure 7.39). The parameter estimates are again initially oscillatory and biased. the comments in case a.3) are applicable here (Figures 7.40-7.42).

7.5) Conclusion

To conclude this chapter we note that both the k-step ahead predictor and the nonlinear GMV controllers coped well with the nonlinearily modelled muscle relaxation system. One source of concern is perhaps the apparent inability of the estimation procedure in the nonlinear GMV case to produce unbiased estimates in the presence of white noise of reasonably small amplitude which could be attributed to the dilemna discussed in case a.3). The consolation in this case is that these biased estimates do not influence the control in a harmful way. As long as the control is good, and all indications are that it will be very good in practice where constant set-points are required, the 'correctness' of the parameter estimates is not very important since inference of useful information as regards relaxant kinetics which is of great importance to pharmacologists is nearly impossible given the complexity of the NARMAX model representations.

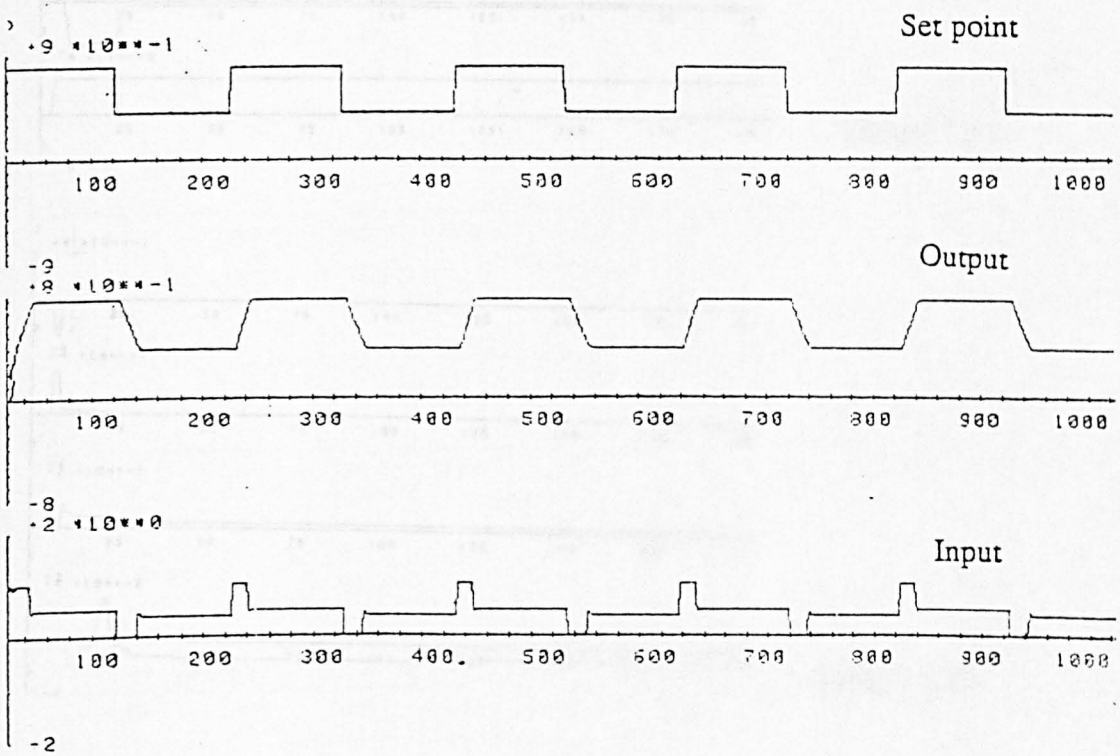


Figure 7.35: Nonlinear GMV controlled system of Atracurium (NOI) with a PI

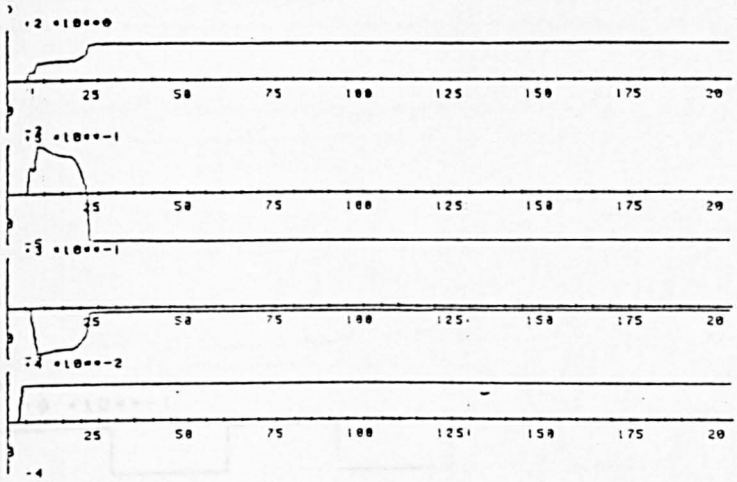


Figure 7.36

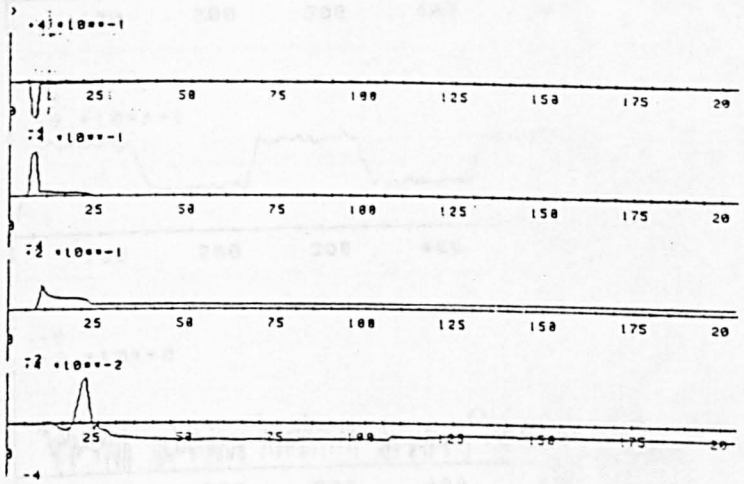


Figure 7.37

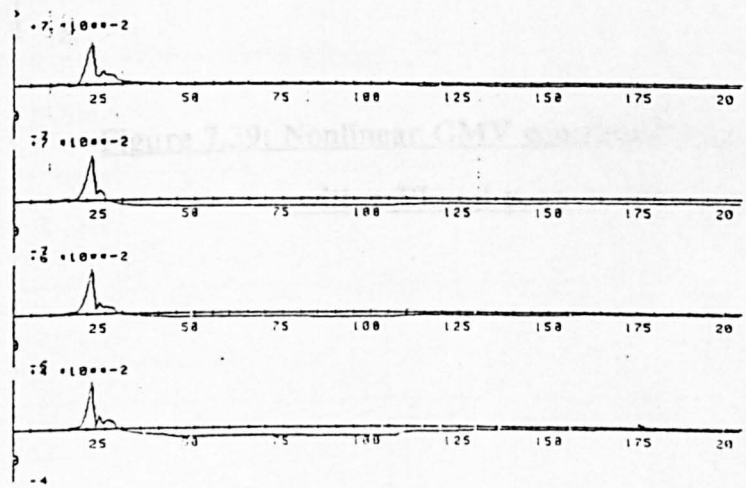


Figure 7.38

Figures 7.36-7.38: Parameter estimates

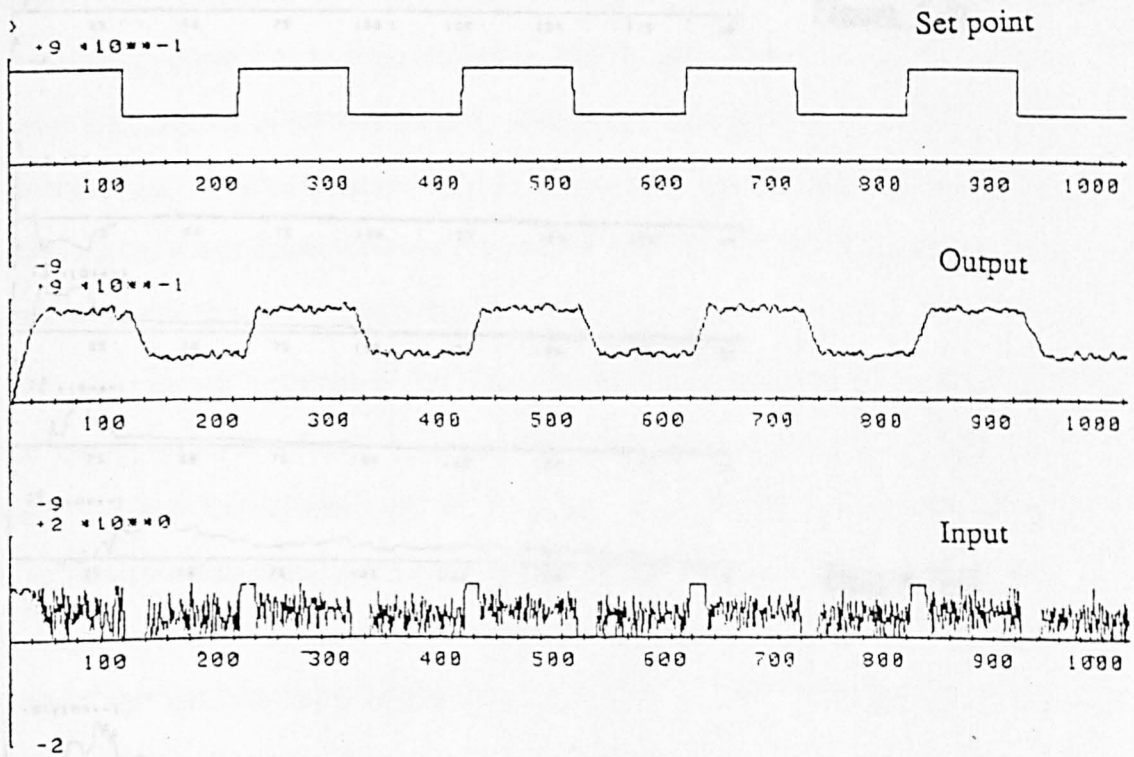


Figure 7.39: Nonlinear GMV controlled system of Atracurium (NLI) with a PI and presence of a white noise

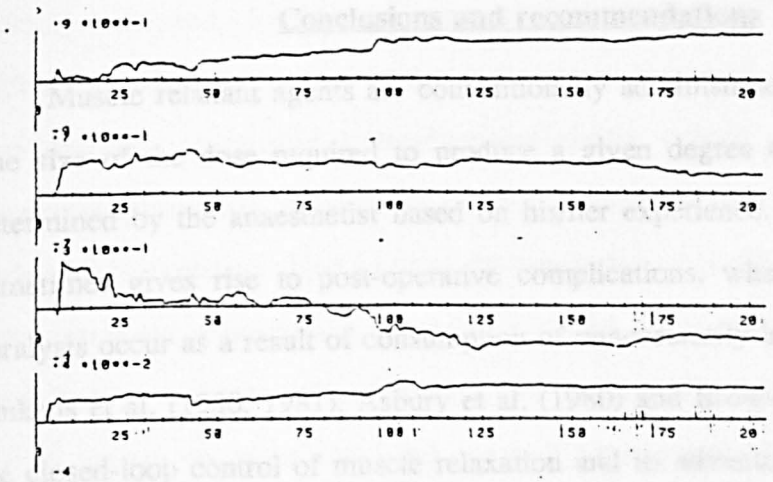


Figure 7.40

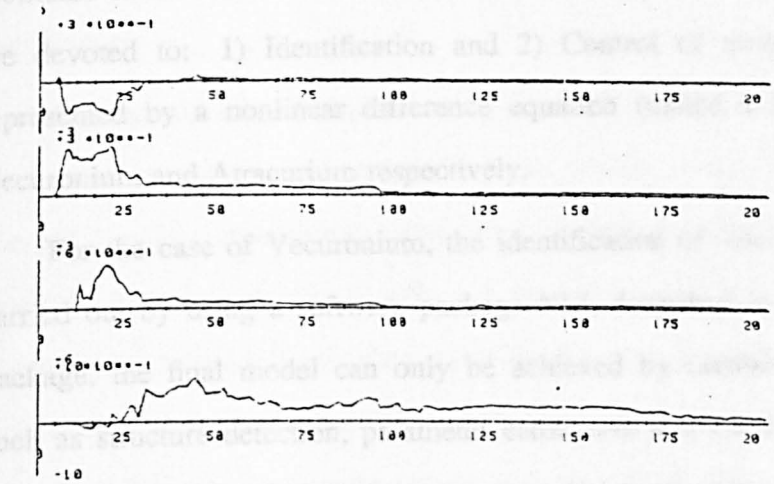


Figure 7.41

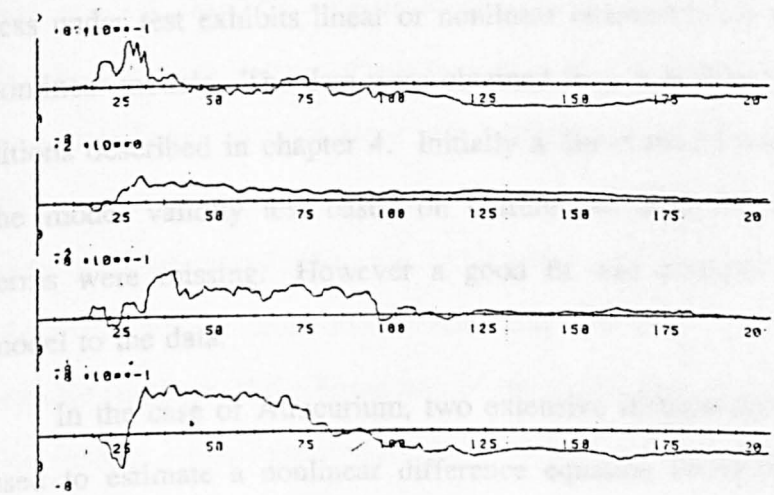


Figure 7.42

Figures 7.40-7.42: Parameter estimates

CHAPTER 8

Conclusions and recommendations

Muscle relaxant agents are conventionally administrated by bolus injections. The size of the dose required to produce a given degree of paralysis is usually determined by the anaesthetist based on his/her experience. This manual method sometimes gives rise to post-operative complications, when conditions of over-paralysis occur as a result of consumption of unnecessarily large amounts of drug. Linkens et al, (1980, 1981), Asbury et al, (1980) and Brown et al , (1980) studied the closed-loop control of muscle relaxation and its advantages in overcoming the problems encountered in manual control. Two major individual parts of this thesis are devoted to: 1) Identification and 2) Control of muscle relaxant dynamics represented by a nonlinear difference equation (called a NARMAX model) for Vecuronium and Atracurium respectively.

For the case of Vecuronium, the identification of the NARMAX model was carried out by using a software package NLI, described in chapter 3. Using this package, the final model can only be achieved by combining several realization such as structure detection, parameter estimation and model validity tests. In the early stages of experimenting on a process it is important to determine if the process under test exhibits linear or nonlinear characteristics which warrant linear or nonlinear models. The data were obtained from a healthy mongrel dog under conditions described in chapter 4. Initially a linear model was fitted to the data, but the model validity test based on correlations analysis indicated that nonlinear terms were missing. However a good fit was obtained by fitting a nonlinear model to the data.

In the case of Atracurium, two extensive identification techniques have been used to estimate a nonlinear difference equation (NARMAX model). The data were obtained under the conditions described in chapter 4. For the first method, a

nonlinear identification package with an extended recursive least squares algorithm (ERLS) was used. For the second method, a nonlinear orthogonal identification package (NOI) consisting of several suites of programs for data generation, structure detection, parameter estimation and model validity tests was used. The parameter estimation program is composed of two algorithms : The extended orthogonal estimation algorithm and the prediction error algorithm. The first algorithm provided a quick way of fitting a model, while the second algorithm was slow but gave improved estimates. The orthogonal property of the algorithm allows parameters to be estimated one at a time by repeated application . Additional terms were added without the need to re-estimate all the previous coefficients.

Pharmacokinetics and pharmacodynamics identification for both muscle relaxants was undertaken in chapter 5. It is widely considered that drug response models should comprise two parts. One represents linear pharmacokinetics (transport of the drug via the blood) and the other represents nonlinear effect pharmacodynamics (which are often modelled by a static characteristic using the Hill equation). The aim was to obtain the parameters for the Hill equation. In the case of the Vecuronium drug values of $\alpha = 2.63$, D=0.5, and for the Atracurium case $\alpha = 3.5$, D=0.62 were obtained. These values are commensurate with those for similar drugs such as Pancuronium. An alternative approach was to use the cross-correlation methods based on Volterra series (Billings and Fakhouri, 1982). These methods require many data points (typically 5000 points) to explore the underlying structure. This was achieved by an off-line simulation using NARMAX models for both muscle relaxants. From the cross-correlation results an impulse response of the linear part of the model was obtained. Using an optimisation routine, a two exponential transfer function was fitted to the time response for Vecuronium, giving

$\tau_1 = 0.96$ min, $\tau_2 = 6.31$ min

and for Atracurium

$\tau_1 = 1.6$ min, $\tau_2 = 22$ min

Also, neither of the cross-correlation conditions showed that the structure was a Hammerstein or a Wiener model. The reason for a Wiener model not being indicated, when this might be expected from physiological knowledge, whereby the linear part can be identified separately, is probably due to the short number of data points and not enough probing in the input used in identifying the models.

The second part of the research consisted of different strategies of control applied to muscle relaxant administration, being three-term PID controller, general minimum variance (Clarke et al, 1975), generalised predictive controller (Clarke and Gawthrop, 1987), k-step ahead predictor and an extension of the GMV self-tuner to nonlinear system modelled by a NARMAX model (Sales,1988). The use of a three-term PID controller showed fairly satisfactory control for both Vecuronium and Atracurium. The responses were obtained with the model parameters fixed to those values for which controller parameter optimization was performed. The general minimum variance approach showed a lack of robustness to control the relaxant dynamics represented by the NARMAX model, whereas the general predictive controller developed by Clarke et al, (1987) demonstrated far superior performance via simulation studies . This method is capable of stable control of processes with variable parameters, with variable dead-time, and with a model order which changes instantaneously provided that the input/output data are sufficiently rich to allow reasonable plant identification. The generalised predictor controller approach proved to be very robust compared to the general minimum variance, with both algorithms viewing the NARMAX model as linear.

A direct digital based k-step ahead predictor and a control weighted self-tuning minimum variance controller with a nonlinear difference equation structure

was also investigated. This method (Sales, 1988) was used on a liquid level system and the results were very encouraging. The same method was applied for both muscle relaxant dynamics (Vecuronium and Atracurium). The research has shown that this approach is superior to the general minimum variance algorithm.
As a result of this research, modern methods of nonlinear systems and identification have been applied successfully to practical biomedical processes. Nonlinear effect are commonplace in the life sciences, and good models for such systems are increasingly important for decision support systems and automated on-line drug infusion approaches. The work reported has shown that interdisciplinary research in these areas may lead to improved systems design and performance.

As suggestions for further work, in the case of Vecuronium and Atracurium, long sets of data with probing inputs are needed. Future research should be extended to the design of adaptive controllers in biological systems by extending the general predictive control and possibly the pole-placement method to the NARMAX model since nonlinearity recurrs frequently in biological systems.
A further suggestion is the development of a package similar to PSICON, but in which the equivalent of the PSI part allows simulation of discrete systems, so as to be able to simulate a NARMAX model in one task and control it from another. This will help in the validation of the controller algorithms in that one can use pseudo- coded ADC/DAC to see the effect of limited accuracy on the system performance. This approach has been used successfully in developing self-tuning algorithms validated in clinical trials on humans.

REFERENCES

-**Aarow, J. Gissen, M.D.** "Clinical of muscle relaxation", Amer. soc of Anesthesiologists.Inc, 1, pp 57-68, 1973.

-**Ali, H.H., and Savarese, J.J.** "Monitoring of neuromuscular function", Anaesthesiology, 45, p 216, 1976.

-**Ali, H.H., Utting, J.E, and Gray, T.C.** "Stimulus frequency in the detection of neuromuscular block in human", Br.J.Anaesth 42, pp 967-072, 1970.

-**Ali, H.H., Wilson, R.S., and Savarese, J.J.** "Effect of tubocurarine on indirectly elicited train-of-four muscle response and respiratory mechanics in conscious human volunteers", Br.J.Anaesth 47, pp 570-574, 1975.

-**Arbib, A.R., and Zeiger, H.P.** "On the relevance of abstract algebra to control theory", Automatics, 5, pp 589-606, 1969.

-**Asbury, A.J., Brown, B.H., and Linkens, D.A.** "Control of neuromuscular blockade by external feedback mechanisms", Br.j.Anaesth., 52, p 633, 1980.

-**Astrom, K.J., and Wittemark, B.** "On self-tunig regulators", Automatica, 9, pp 185-199, 1973.

-**Batey, D.J., Sterling, M.J.H., Antcliffe, D.J., and Billings, S.A.** "The design and implementation of an interactive data analysis package for a process computer", Computer aided design, 7, 4, pp 225-269, 1975.

-**Billings, S.A.** "Identification of nonlinear systems-a survey" , Proc. IEE, 127, pp 272-285, 1980.

-**Billings, S.A.** "An overview of nonlinear systems dentification", Department of Control Engineering University of Sheffield. Research report 271, january 1985.

-**Billings, S.A.** "Introduction to nonlinear systems analysis and identification", SERC Vac, School on ' Sgnal processing for control ', University of Warwick, pp 15-20, sept, 1985.

-**Billings, S.A., and Fadzil, M.B.** "The practical identification of systems with nonlinearities", 7th IFAC. Symp. Ident and Syst. Par. Est York 1985.

-**Billings, S.A., and Fadzil, M.B.** "Identification of a nonlinear difference equation model of an industrial diesel generator", Research report, Dept of control Eng., University of Sheffield, 1984.

-**Billings, S.A., and Fakhouri, S.Y.** "Identification of nonlinear systems using the Wiener model", Eletronics letters, 13, 17, pp 502-504, 1977.

-**Billings, S.A., and Fakhouri, S.Y.** "Theory of separable process with applications to the identification of nonlinear system", Proc. IEE, 125, pp 1051-1057, 1978.

-**Billings, S.A., and Fakhouri, S.Y.** "Identification of systems containing linear dynamics and static nonlinear elements", Automatica, 18, pp 15-26, 1982.

-**Billings, S.A., Korenberg, M., and Liu, Y.P.** "An orthogonal parameters estimation algorithm for nonlinear stochastic systems", Research report 307, Sheffield University, March 1987.

-**Billings, S.A., and Leontarities, I.J.** "Identification of nonlinear systems using parameter estimation techniques", Proc. IEE. Conference, control and its applications University of Warwick, pp 183-187, 1981.

-**Billings, S.A., and Leontarities, I.J.** "Parameter estimation techniques for nonlinear systems", 6th IFAC Symp. Ident and Syt.Par.Est, Washington, D.C, 1982.

-**Billings, S.A., and Stang, K.M.** "User manual for the fast nonlinear orthogonal identification package", Department of Control Engineering University of Sheffield, 1987.

-**Billings, S.A., and Sterling, M.J.H.** "SPAID- A user guide", Department of Control Engineering University of Sheffield, 1979.

-**Billings, S.A., and Voon, W.S.F.** "Structure detection tests and model validity tests in the identification of nonlinear systems", Proc. IEE, part D, 130, pp 193-199, 1983.

-**Billings, S.A., and Voon, W.S.F.** "Least squares parameter estimation algorithms for nonlinear systems", Int. J. Syst. Sci, 15, 6, pp 601-615, 1984.

-**Billings, S.A., and Voon, W.S.F.** "Correlation based on model validity tests for nonlinear models", Int. J. Control, 44, pp 235-244, 1986b.

-**Birks, R.H.E., Huxley., Katz, B.** "The fine structure of the neuromuscular junction of the frog", J. Phy, 150, pp 134-144, 1960.

-**Brown, B.H., Asbury, A.J., Linkens, D.A., Perks, R., and Anthony, M.** "Closed loop control of muscle relaxation during surgery", Clin, Phys and Physiol. Meas, 1, p 203, 1980.

-**Brown, R.F., and Godfrey, K.R.** "Problems of determinacy in compartmental modelling with application to bilirubin kinetics", Math, Biosci, 40, pp 205-224, 1978.

-**Cass, N.H., Lampard, D.G., Brown, W.A., and Coles, J.R.** "Computer controlled muscle relaxation: A comparison of four muscle relaxants in the sheep", Anaesth. and intensive care, IV, 4, pp 16-22, 1976.

-**Clarke, D.W., Cope, S.N., and Gawthrop, P.S.** "Feasibility study of the application of microprocessor to self-tuning controllers", Tech. Rep. 1137/75. Department of engineering of sciences, University of Oxford, 1975.

-**Clarke, D.W., and Gawthrop, P.S.** "Self-tuning controller" Proc. IEE, 122, pp 929-934, 1975.

-**Clarke, D.W., and Gawthrop, P.S.** "Self-tuning controller", Proc. IEE, 126, 6, june, pp 633-640, 1979.

-**Clarke, D.W., Mohtadi ,C., and Tuffs, P.S.** "Generalised predictive control PART I. The basic algorithm". Automatica, 23, 2, pp 137-148, 1987

-**Clarke, D.W., Mohtadi, C., and Tuffs, P.S.** "Generalised predictive control PART II. Extention and interpretation", Automatica, 23, 2, pp 149-160, 1987.

-**Couteaux, R.** "Nouvelles observations sur la structure de la plaque motrice et interpretation des rapports myo-neuraux", C.R.Soc, Biol, 138, 976, 1944.

-**De Keyser, R.M.C., and Van Canwenberghe, A.R.** "Typical application possibilities for self-tuning predictive control", IFAC Symp Ident. Syst.Par.Est. Washington, 1982.

-**Denai, M.** "Self-adaptive PID control of drug induced muscle relaxation during anaesthesia", PhD thesis, Sheffield University department of control engineering, 1988.

-**Epstein, R.M., and Epstein, R.A.** "The electromyograph and the mechanical response of indirectly stimulated muscle in anaesthetised man following curarization", Anaesthesiology, 38, p 21, 1973.

-**Epstein, R.M., and Epstein, R.A.** "Electromyography in evaluation of the response to muscle relaxants", In.R.L. Katz (Ed.) Muscle relaxants. Amsterdam: Excerpta Medica, 1975.

-**Fadzil, M.B.** "The practical identification of nonlinear systems", PhD thesis Sheffield University department of control engineering, 1985.

-**Fakhouri, S.Y.** "Identification of Volterra kernels of nonlinear systems", Proc IEE, Part D, 127, pp 246-304, 1980.

-**Feldman, S.A.** "Drugs in anaesthesia: Mechanics of action" ,1987.

-**Fleischli, G., and Cohen, E.N.** "An analogue computer simulation for the distribution of D-tubocurarine", Anaesthesiology, 27, p 64, 1966.

-**Gawthrop, P.J.** "Self-tuning PID controller. Algorithm and implementation", IEEE Trans, Auto.control, A-C, 31, 3, p 201, 1986.

-**Gibaldi , M., Levy, G., and Hayton, W.** "Kinetics of the elimination and Neuromuscular blocking effect of d-Tubocurarine in man", Anesthesiology, 36, pp 213-218, 225, 1972.

-**Gibaldi, M., and Perrier, D.** "Pharmacokinetics", (Marcel Dekker, Inc, 1975).

-**Goodwin, G.C., and Payne, R.L.** "Dynamic system identification - Experiment design and data analysis", Academic Press, 1977.

-**Goodwin, G.C., and Sink, S.** "Adaptive filtering, prediction and control", Prentice Hall, 1984.

-**Graham, J., and Gerard, R.W.** "Membrane potentials and excitation of Impaled single muscle fibres", J. Cell. and comp. Physiol, 28, 99, 1946.

-**Gutman, E., and Young, J.Z.** "The re-innerration of muscle after various periods of Atrophy", J.Anat, 78, 15, 1944.

-**Haber, R., and Keviczky, L.** "Identification of nonlinear dynamic systems", 4th IFAC Symp.Ident and Syst. Par. Est. Tblissi, 1976.

-**Ham, J., Miller, R.D., Sheiner, L.B., and Matteo, R.S.** "Dosage schedule independenceof d-Tubocurarine pharmacokinetics and pharmacodynamics and recovery of neuromuscular function ", Anesthesiology, 50, pp 528-533, 1979.

-**Hug, C.C., and Roberts, C.P.** "Pharmacokinetics of anaesthesia", 1984.

-**Hull, C.J., English, M.J.M., and Sibbald, A.** "Fazadinium and Pancuronium: a pharmacodynamic study", Br.J. Anaesth, 52, pp 1209-1220, 1980.

-**Hull, C.J., Van Beem, H.B.H., Mcleod, K., Sibbald, A., and Watson, M.J.** "A pharmacodynamic model for Pancuronuim", Br.J.Anaesth, 50, p 113, 1978.

-**Iserman, R.** "Digital control systems", Springer Verlay, 1972.

-**Jacquez, J.A.** "Compartmental analysis in biology and medicine", (Elsevier Publishing Co, 1972).

-**Katz, R.L.** "What's new and old but true in muscle relaxants", The Amer, soc of anaesth, 2, pp 117-132, 1974.

-**Kim, J.H and Choi, K.K.** "Self-tuning discrete PID controllers", IEEE Trans.Ind.Electronics, I-E 34, 2, 1987.

-**Korenberg, M.** "Statistical identification of parallel cascades of linear and non-linear systems", IFAC Symp. Ident., and Sys.Par.Est., Arlington, Virginia, pp 580-585, 1982.

-**Kraus, T.W and Myron, T.J.** "Self-tuning PID controller uses pattern recognition approach", Control-Eng, 31, 6, p 106, 1984.

-**Leontaritis, I.J., and Billings, S.A.** "Input-output parametric models for non-linear systems.", PartI. Deterministic nonlinear systems. PartII Stocastic nonlinear systems, Int. J. Control, 41, pp 303-344, 1985a,b.

-**Linkens, D.A.** "Biological systems, modelling and control", 1979.

-**Linkens, D.A.** "Control of muscle relaxation during surgery", Biomed, Meas, Infor Contr, 1, 1, pp 31-39, 1986.

-**Linkens, D.A., and Asbury, A.J.** "Non-invasive measurement of Vecuronium pharmacokinetics and pharmacodynamics using systems identification techniques", Br.J. Anaesth, 57, 829, 1985.

-**Linkens, D.A., Asbury, A.J., and Brown, B.H.** "On-line control relaxant administration during anaesthesia", IEE conference on control and its application, Warwick, March, 1981.

-**Linkens, D.A., Asbury, A.J., Rimmer, S.J., and Menad, M.** "Self-tuning control of muscle relaxation during anaesthesia", IEEE conference on applications of adaptive and multivariable control. Hull, pp 96-102, 1982.

-**Linkens, D.A., and Mahfouf, M.** "Fuzzy logic knowledge based control for muscle relaxant anaesthesia", IFAC, Symposium, modelling and control in biomedical system. Pre-prints of IFAC, pp 123-128, 1988.

-**Linkens, D.A., and Mahfouf, M.** "Generilzed predictive control of muscle relaxant anaesthesia", Worshop on Decision support for patient management: Measurement, modelling and control. City university London, 1989.

-**Linkens, D.A., Menad, M., Mort, N., Gray, L.S., and Bennett, S.** "PSICON - A simulation package for the design of digitally controlled systems", Trans. Inst. MC, 4, 3, pp 153, 1982.

-**Linkens, D.A., Menad, M., and Rimmer, S.J.** "Identification and control of muscle relaxant anaesthesia", Proc IEE. Pt. D. 129, pp 136-141, 1982.

-**Macgregor, J.F., Wright, J.D., and Hong, H.N.** "Optimal tuning of digital PID controllers using dynamic stochastic models", IEC Process Des. Dev. 14, pp 398-402, 1975.

-**Marmarellis, P.Z., and Marmarellis, V.Z.** "Analysis of physiological systems the white noise approach", Plenum Press. 1978.

-**Menad, M.** "Feedback control of drug administration for muscle relaxation", PhD thesis, University of Sheffield, 1984.

-**Narendra, K.S., and Gallman, P.** "An iterative method for the identification of nonlinear systems using Hammerstein models", IEEE trans Auto, Control., AC-11, pp 546-550, 1966.

-**Nelder and Mead.** "A simplex method for function minimization", Complet. J, 7, p 308, 1965.

-**Paton, W.D.M., and Waud, D.R.** "The margin of safety of neuromuscular transmission", J. Physiol (Lond), 191, 59, 1967.

-**Pugachev, V.S.** "Theory of random functions and its application to control problems", 5, 1962.

-**Rametti, L.B.** "On-line control of d-Tubocurarine induced muscle relaxation", Ph.D thesis University of Cape Town, 1985.

-**Raven, H.R.** "Automatic control engineering", 1978.

-**Rike, W.F.** "Prejunctional effects of neuromuscular blocking and facilitatory drugs in muscle relaxants", Amsterdam: Excerpta Medica, p 59, 1975.

-**Sales, R.K.** "Self-tuning control of nonlinear systems", PhD thesis, University of Sheffield, 1988.

-**Savarese, J.J., and Kitz, R.J.** "Pharmacology of relaxants", The Amer, Soc, of Anaesth, mc, 3, pp 153-170, 1975.

-**Schetzen, M.** "The Volterra and Wiener theories of nonlinear systems", Wiley. 1980.

-**Shanks, C.A., Somogyi, A.A., Ramzan, M.I., and Triggs, E.J.** "Tubocurarine and Pancuronium: A pharmacokinetics view", Anaesth. intens care 8, 4, 1980.

-**Sheiner, L.B., Stanski, D.R., Vozeh, S., and al.** "Simultaneous modelling of pharmacokinetics and pharmacodynamics: Application to d-Tubocurarine", Clin. Pharmacol, Ther, 25, p 358-371, 1979.

-**Sheppard, L.C., Shotts, J.F., Roberson, N.F., Wallace, F.D., and Kouchoukos, N.T.** "Computer controlled infusion of vasoactive drugs in post-cardiac surgical patients", Proc. IsT. Annual conference IEEE Eng. in Medecine and Biology Soc., Denver, Colorado, p 280, 1979.

-**Soderstrom, T. Ljung, L., and Gustavsson, I.** "A theoretical analysis of recursive identification methods", Automatica, 14, pp 231-244, 1978.

-**Stanski, D.R., Ham, S., Miller, R.D., et al.** "pharmacokinetics and pharmacodynamics of d-Tubocurarine during nitrous oxide-narcotic and halothane anaesthesia in man", Anaesthesiology, 5, pp 235-241, 1979.

-**Sumner, D.J., Russell, A.J., and Whiting, B.** "Digoxin kinetics: Multicompartmental analysis and its clinical implications", Br.J. of Clin Pharmacology, 3, pp 221-229, 1976.

-**Van Den Bosch, PPJ.** "PSI-An extended interacrive block orientated simulation program", IFAC Symposium on computer aided design of control systems, Zurich, pp 223-228, 1979.

-**Volterra, V.** "Theory of functionals", Blakie and Sons, 1930.

-**Wagner, J.C.** "Kinetics of pharmacological response", J. Theory. Biol, 20, pp 173-201, 1968.

-**Wagner, J.C.** "Linear pharmacokinetics equations allowing direct correlation of many needed pharmacokinetic parameters from the coefficients and the exponents of poly-exponential equations which have been fitted to the data", J . pharmacokinetic. Bioparm Li, pp 443-467, 1976.

-**Waud, D.R.** "The nature of depolarization block", Anaesthesiology, 29, 1014, 1968.

-**Waud, D.R., and Waud, B.E.** "The relation between tetanic fade and receptor occlusion in the presence of competitive neuromuscular block", Anaesthsiology, 35, 456, 1971.

-**Weatherley, B.C., Williams, S.G., and Neil, E.A.M.** "pharmacokinetics, pharmacodynamics and dose response relationship of atracurium administered IV", Br.J.Anaesth, 55, pp 395, 1983.

-**Wellstead, P.E., and Zanker, P.M.** "The techniques of self-tuning", UMIST control System Center report 432, 1978.

-**Wellstead, P.E., Prager, D.L., and Zanker, P.M.** "Pole assignment self-tuning regulator", PROC.IEE, 126, 8, 781-787, 1979.

-**Whiting, B., and Kelman, A.W.** "The modelling of the drug response", Clinical Science, 59, pp 311-315, 1980.

-**Woobbury, J.W., Gordon, A.M., and Conrad, J.T.** " Muscle ", Physiology and Biophysics-Ruch-Patton, pp 113-152, 1965.

-**Young, P.C.** "An instrumental variable method for real time identification of a noisy process", Automatica, 6, pp 271-287, 1970.

-**Ziegler, J.G., and Nicholas, N.B.** "Optimum settings for automatic controllers", Trans. Amer. Soc. Mech. Eng, 64, pp 759, 1942.

APPENDIX A

System representation

Consider a causul discrete time-invariant system with input set U, output set Y and define the input-output or response map as the map that gives the system output y(t) for $t \geq 0$ as a function of the input sequence in the time interval from 0 to t.

$$y(t) = f^t[u(t),u(t-1), \cdots ,u(0)] \tag{A.1}$$

where

$f^t[.]$ indicates the time interval of the input (Arbib and Zeiger, 1969). The method basically consists of grouping together all the input sequences for which the response of the input-output map after these sequences are applied is the same. For example, define two input sequences as

$$w_1 = (u(t_1),u(t_1 - 1), \cdots ,u(0))$$
$$w_2 = (u(t_2),u(t_2 - 2), \cdots ,u(0))$$

The sequenses w_1 and w_2 are said to be equivalent if and only if

$$f^{t_1+1}(u_{1,1},w_1) = f^{t_2+1}(u_{1,1},w_2) \quad \textit{for all} \quad u_{1,1} \in U$$
$$f^{t_1+2}(u_{2,2},u_{2,1},w_1) = f^{t_2+2}(u_{2,2},u_{2,1},w_2) \quad \textit{for all } u_{2,1},\ u_{2,2} \in U$$
$$f^{t_1+3}(u_{3,3},u_{3,2},u_{3,1},w_1) = f^{t_2+3}(u_{3,3},u_{3,2},u_{3,1},w_2)$$
$$\textit{for all } u_{3,1},\ u_{3,2},\ u_{3,3} \in U \quad \textit{etc.} \tag{A.2}$$

The above equivalence relationship is the state set of the Nerode realization. (Billings and Leontaritis, 1984) detailed the system representation. Here a brief summary is given

An input-output map is thus called finite time observable if and only if there exists an integer k such that for any two input sequences w_1 and w_2

$$f^{t_1+1}(u_{1,1},w_1) = f^{t_2+1}(u_{1,1},w_2) \quad \textit{for all } u_{1,1} \in U$$
$$f^{t_1+2}(u_{2,2},u_{2,1},w_1) = f^{t_2+2}(u_{2,2},u_{2,1},w_2) \quad \textit{for all } u_{2,1},\ u_{2,2} \in U$$
$$f^{t_1+k}(u_{k,k},u_{k,k-1}, \cdots ,u_{k,1},w_1) = f^{t_2+k}(u_{k,k},u_{k,k-1}, \cdots ,u_{k,1},w_2)$$
$$\textit{for all } u_{k,k}, \cdots ,u_{k,1} \in U$$

implies that

$$f^{t_1+k+1}(u_{k+1,k+1}, \cdots, u_{k+1,1}w_1) = f^{t_2+k+1}(u_{k+1,k+1}, \cdots, u_{k+1,1}w_2)$$
$$\text{for all } u_{k+1,k+1} \cdots \varepsilon\, U$$

and

$$f^{t_1+k+2}(u_{k+2,k+2}, \cdots, u_{k+2,1}w_1) = f^{t_2+k+2}(u_{k+2,k+2} \cdots w_2)$$
$$\text{for all } u_{k+2,k+2} \cdots \varepsilon\, U \text{ etc.} \qquad \text{(A.3)}$$

The equivalence conditions of equation (A.2) can be replaced by the conditions of equation (A.3).

It would however be far more realistic if the response to a single input sequence were sufficient to determine the equivalence of two inputs sequences w_1 and w_2 in equation (2) or (3). If the conditions in equation (A.2) or (A.3) were valid for only one $u_1 \varepsilon U$, one $u_2 \varepsilon U$, where $u_1 = u_{1,1} = u_{1,2} = u_{1,3}$.... ; $u_2 = u_{2,1} = u_{2,2} \cdots$ they would be valid for any $u_{1,1}, u_{1,2}, u_{1,3}$ etc. If in addition, $u_1, u_2 \cdots$ could be any member of the set U, then the input/output map will be called input-independent single input observable. Hence it can be proved that there exists a function F(.) such that

$$f^{t_1+k+1}(u_{k+1}, u_k, \cdots, u_1, w_1) = F[u_{k+1}, u_k, \cdots, u_1 f^{t_1+1}(u_1, w_1),$$
$$f^{t_1+2}(u_2, u_1, w_1, \cdots f^{t_1+k}(u_k, u_{k-1}....u_1, w_1)] \qquad \text{(A.4)}$$

Which can be expressed as

$$y(t_1+k+1) = F[y(t_1+k), \cdots, y(t_1+1), u(t_1+k+1), u(t_1+k) \cdots u(t_1+1)]$$

or substituting $t = t_1+k+1$

$$y(t) = F[y(t-1), \cdots, y(t-k), u(t), \cdots, u(t-k)] \qquad \text{(A.5)}$$

This model represents a generalisation of the well known ARMA model to non-linear system. When the noise terms are included the model is called a nonlinear ARMAX or NARMAX model.

A similar derivation can be followed for stochastic systems. It can be shown that

$$y(t) = F^*[y(t-1), \cdots, y(t-k), u(t), \cdots, u(t-k), \varepsilon(t-1)$$
$$, \cdots, \varepsilon(t-k)] + \varepsilon(t) \qquad \text{(A.6)}$$

where

$\varepsilon(t)$ is the prediction error and it is independent from y^{t-1},u^t

and

$$E[\varepsilon(t)|y^{t-1},u^t] = 0$$

and

$$\varepsilon(t) = y(t) - \hat{y}(t) = y(t) - f(y^{t-1},u^t)$$

where

$$\hat{y}(t) = E[y(t)|y^{t-1},u^t] = f(y^{t-1},u^t)$$

The future value $\varepsilon(t+1)$ will be unpredictable from ε^t .

APPENDIX B

The orthogonal least squares algorithm

Parameter estimation in NOI is based on the orthogonal least squares algorithm. This algorithm was derived for stochastic nonlinear systems which can be represented by a NARMAX model (Billings, Korenberg and Liu, 1987). The advantages of this algorithm lies in its superior efficiency and that it provides information regarding which terms in the model are significant. This is often vital in the identification of NARMAX models.

3.1) NARMAX model

A wide class of nonlinear systems can be represented by the following expression (Leontaritis and Billings, 1985 a, b")

$$y(t) = F^l[y(t-1), \cdots, y(t-N_y), u(t), \cdots, u(t-N_u), \varepsilon(t-1), \cdots, \varepsilon(t-N_\varepsilon)] + \varepsilon(t) \tag{B.1}$$

Where

u(t) and y(t) represent the measured input and output respectively.

N_u, N_y, N_ε represent the number of the lags in the input, output and the predictor error respectively.

Where

$\varepsilon(t)$ is the predictor error defined as

$$\varepsilon = y(t) - \hat{y}(t) \tag{B.2}$$

and

$E[\varepsilon(t)|y^{t-1}, u^t] = 0$

$y^{t-1} = [y(t-1), y(t-2), \cdots, y(1)]^T$

$u^t = [u(t), u(t-1), \cdots, u(1)]^T$

$F^l[.]$ is some nonlinear function.

A time delay in the input and dc level can easily be accommodated and equation (B.1) becomes

$$y(t) = dc + F^l[y(t-1), \cdots, y(t-N_y), u(t), \cdots, u(t-N_u), \varepsilon(t-1), \cdots, \varepsilon(t-N_\varepsilon)] + \varepsilon(t) \quad \text{(B.3)}$$

Expanding equation (B.3) by defining the function $F^l[.]$ to be a polynomial of degree l gives

$$y(t) = \sum_{m=0}^{M} \theta_m p_m(t) + \varepsilon(t) \quad \text{(B.4)}$$

Direct least squares estimation on equation (B.4) may involve an excessive number of terms. An orthogonal least squares algorithm (Korenberg, Billings and Liu, 1987) can overcome these difficulties.

1.3) Orthogonal regression

The Orthogonal least squares is much more efficient than the least squares algorithm. The advantages of the algorithm lie mainly in the fact that it provides information regarding which terms in the model are significant. (Korenberg, Billings and Liu, 1987) detailed the derivation of this algorithm, only the main results are quoted here.

Consider a linear-in-parameters expression of equation (B.4)

$$Y(t) = \sum_{i=1}^{M} w_i(t) g_i + \zeta(t) \quad \text{(B.5)}$$

where

$$w_i(t) = p_1(t) \quad \text{(B.6)}$$

$$w_k(t) = p_k(t) - \sum_{i=1}^{k-1} \alpha_{ik} w_i(t), \quad k = 2, \cdots, M \quad \text{(B.7)}$$

and

$$\alpha_{ij} = \frac{\overline{w_i(t) p_j(t)}}{\overline{w_i^2(t)}}, \quad i<j, \quad j = 1,2, \cdots, M \quad \text{(B.8)}$$

Here, the bar $\overline{\ }$ denotes time averaging. The estimates of the coefficients g_i are given by

$$\hat{g}_i = \frac{\overline{w_i(t) Y(t)}}{\overline{w_i^2(t)}}, \quad i = 1, \cdots, M \quad \text{(B.9)}$$

The coefficients of the original equation can easly be obtained according to the formula

$$\hat{\theta}_M = \hat{g}_M$$
$$\hat{\theta}_i = \hat{g}_i - \sum_{j=i+1}^{M} \alpha_{ij}\hat{\theta}_i, \quad i = M-1, M-2, \cdots, 1 \tag{B.10}$$

The estimates (equation (B.10)) are unbiased and the standard deviations of the estimates are given by

$$std(\hat{\theta}_i) = \sigma\sqrt{\frac{1}{N}\sum_{j=1}^{M}\frac{t_{ij}^2}{\overline{w_j^2(t)}}}, \quad i = 1, \cdots, M \tag{B.11}$$

where

$$t_{ij} = \begin{cases} 1 & i = j \\ -\sum_{k=i+1}^{j} \alpha_{ik}t_{kj}, & 0 < i < j, \ j = 1, \cdots, M \end{cases} \tag{B.12}$$

and $\sigma^2 = E[\zeta^2(t)]$. An estimate of σ can be obtained from

$$\sigma = \sqrt{\frac{1}{N-M}\sum_{t=1}^{N}\left(Y(t) - \sum_{i=1}^{M} w_i(t)\hat{g}_i\right)^2} \tag{B.13}$$

Because the orthgonality

$$\overline{Y^2(t)} = \sum_{i=1}^{M} g_i^2\overline{w_i^2(t)} + \sigma^2 \tag{B.14}$$

The ratio of the reduction in the sum of squared errors due the i'th term is therfore given by

$$[rer]_i = \frac{g_i^2\overline{w_i^2(t)}}{\overline{Y^2(t)}}, \quad i = 1, \cdots, M \tag{B.15}$$

The above equation can be used as a criterion for determining the significant terms in the model and insignificant terms may be removed from the orthogonal equation. This is equivalent to removing corresponding $p_i(t)$ ' s from the original model leading to a less complex model.

Interpreting equation (B.15) requires care, since if the position of $p_i(t)$ is altered in equation (B.4) a different value for $[rer]_i$ is given. In general, the earlier $p_i(t)$

appears in equation (B.4) the larger $[rer]_i$ will be. In an attempt to overcome this difficulty an orthogonal regression procedure based on the idea of stepwise regression is suggested.

In the initial stage, all $p_i(t)$, $i = 1,..,M$ are considered as possible candidates for $w_i(t)$. For i=1,..,M, calculate

$$w_1^{(i)}(t) = p_i(t)$$

$$\hat{g}_1^{(i)} = \frac{\overline{w_1^{(i)}Y(t)}}{\overline{(w_1^{(i)}(t))^2}}$$

$$[rer]_1^{(i)} = \frac{(g_1^{(i)})^2\overline{(w_1^{(i)}(t))^2}}{\overline{Y^2(t)}}$$

Find the maximum of $[rer]_1^{(i)}$, say $[rer]_1^{(i)} = \max[rer]_1^{(i)}$, $1 <= i <= M$. Then the one-term orthogonal equation is selected

$$Y(t) = w_1(t)g_1 + \zeta(t)$$

with

$$w_1(t) = w_1^{(j)}(t),\ \hat{g}_1 = \hat{g}_1^{(j)} \text{ and } [rer]_1 = [rer]_1^{(j)}$$

In the second stage, all $p_i(t)$, $i = 1, \cdots, M, i \neq j$ are considered as possible candidates for $w_2(t)$. For $i = 1, \cdots, M$, $i \neq j$, calculate

$$w_2^{(i)}(t) = p_i(t) - \alpha_{12}^{(i)}w_1(t)$$

$$\hat{g}_2^{(i)} = \frac{\overline{w_2^{(i)}(t)Y(t)}}{\overline{(w_2^{(i)}(t))^2}}$$

$$[rer]_2^{(i)} = \frac{(g_2^{(i)})^2\overline{(w_2^{(i)}(t))^2}}{\overline{Y(t)^2}}$$

where

$$\alpha_{12}^{(i)} = \frac{\overline{w_1 p_i(t)}}{\overline{w_1^2(t)}}$$

Find the maximum of $[rer]_2^{(i)}$, say $[rer]_2^{(k)} = \max[rer]_2^{(i)}$, $1<=i<=M$. Then the two-term orthogonal equation is selected

$$Y(t) = w_1(t)g_1 + w_2(t)g_2 + \zeta(t)$$

with

$w_2(t) = w_2^{(k)}(t) = p_k(t) - \alpha_{12} w_1(t)$, $\hat{g}_2 = \hat{g}_2^{(k)}$ and $[rer]_2 = [rer]_2^{(k)}$,

where

$\alpha_{12} = \alpha_{12}^{(k)}$.

The procedure is terminated at the m'th stage when

$1 - \sum_{i=1}^{m} [rer]_i <$ a desired tolerance $n < M$. The final orthogonal equation is selected as

$$Y(t) = \sum_{i=1}^{m} w_i(t) g_i + \zeta(t)$$

Using equations (B.9) and (B.10), it is a straightforward task to calculate the corresponding parameters θ_i in the model containing only m significant terms.

Because of the orthogonality, before adding the i+1'th term to the orthogonal equation, there is no need to remove the effect of the i'th term. In step wise regression a term that was significant at an earlier stage may become insignificant after several other terms are included in the model. To overcome this difficulty, before adding a new term, those terms that are already in the model are tested for their significance using some statistical test (Billings and Voon, 1986b) and if necessary deleted. Such precaution is not needed in the orthogonal regression.

APPENDIX C

LINEARIZATION

Most methods of system analysis have been developed for linear control systems. For a linear control system all the relationships between the variables are linear differential equations, usually with constant coefficients.

Actual control systems usually contain some nonlinear elements. Such elements would in turn yield nonlinear differential equations for the system.

In the following it is shown how the equations for nonlinear elements may be linearized (Raven,1978)

Consider the following equation

$$Y = X^2 \tag{C.1}$$

The plot of equation C.1 is shown below in Figure C.1.1

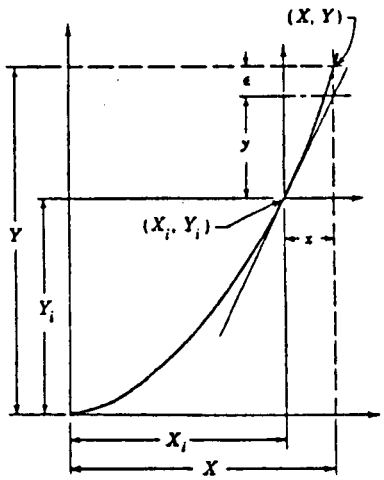


Figure C.1.1

In the vicinity of the point (X_i,Y_i) , the function is closely approximated by the tangent. For example, consider the pair (X,Y) on the curve of the nonlinear function. X is displaced a distance x from X_i. X intersects the nonlinear function a vertical distance $y + \varepsilon$ from Y and it intersects the tangent a distance y from Y_i . Hence the equation for Y is

$$Y = Y_i + y + \varepsilon \approx Y_i + y \tag{C.2}$$

The slope at point (X_i,Y_i) is

$$\frac{y}{x} = \left[\frac{dY}{dX}\right]_i \tag{C.3}$$

or

$$y = \left[\frac{dY}{dX}\right]_i x = \left[\frac{dX^2}{dX}\right]_i x = 2X_i x \tag{C.4}$$

equation C.2 becomes

$$Y \simeq Y_i + 2X_i x \tag{C.5}$$

A general procedure for obtaining a linear approximation is to use the variation ΔY for a function $Y = Y(X_1,X_2, \cdots ,X_n)$ of n independent variables, that is

$$\Delta Y = \left[\frac{\partial Y}{\partial X_1}\right]_i \Delta X_1 + \left[\frac{\partial Y}{\partial X_2}\right]_i \Delta X_2 + \cdots + \left[\frac{\partial Y}{\partial X_n}\right]_i \Delta X_n \tag{C.6}$$

Therefore

$$\begin{bmatrix} y = \Delta Y = & Y - Y_i \\ x_1 = \Delta X_1 = & X_1 - X_{1i} \\ x_2 = \Delta X_2 = & X_2 - X_{2i} \\ \vdots & \vdots \\ x_n = \Delta X_n = & X_n - X_{ni} \end{bmatrix}$$

Thus, the general expression for obtaining a linear approximation for the linear function

$$y = C_1x_1 + C_2x_2 + \cdots + C_nx_n \tag{C.7}$$

Where

$$\left[C_1 = \frac{\partial Y}{\partial X_1}\right]_i, \left[C_2 = \frac{\partial Y}{\partial X_2}\right]_i \text{ etc}$$

The above method is applicable to linearize the NARMAX model represented by equation C.8

$$\begin{aligned} z(t) = {} & k_1z(t-1) + k_2z(t-2) + k_3z(t-3) + k_4z(t-4) \\ & + k_5z(t-5) + k_6u(t-4) + k_7u(t-3)z(t-1) + k_8u(t-4)z(t-1) \\ & + k_9u(t-5)z(t-1) + k_{10}z(t-4)z(t-5) + k_{11}u(t-3)z(t-5) \end{aligned} \tag{C.8}$$

where

k_1 =1.376 k_5 =-0.1750 k_9 =0.004134

k_2 =-0.244 k_6 =0.007367 k_{10} =-0.03932

k_3 =-0.3677 k_7 =0.03702 k_{11} =0.02160

k_4 =0.4154 k_8 =-0.004919

The partial derivatives are:

$$\frac{\partial z(t)}{\partial z(t-1)} = (k_1 + k_7\bar{u}(t-3) + k_8\bar{u}(t-4) + k_9\bar{u}(t-5))(z(t-1) - \bar{z})$$
$$= f_1(z(t-1) - \bar{z})$$
$$\frac{\partial z(t)}{\partial z(t-2)} = (k_2)(z(t-2) - \bar{z})$$
$$= f_2(z(t-2) - \bar{z})$$
$$\frac{\partial z(t)}{\partial z(t-3)} = (k_3)(z(t-3) - \bar{z})$$
$$= f_3(z(t-3) - \bar{z})$$
$$\frac{\partial z(t)}{\partial z(t-4)} = (k_4 + k_{10}\bar{z}(t-5))(z(t-4) - \bar{z})$$
$$= f_4(z(t-4) - \bar{z})$$
$$\frac{\partial z(t)}{\partial z(t-5)} = (k_5 + k_{10}\bar{z}(t-4) + k_{11}\bar{u}(t-3))(z(t-5) - \bar{z})$$
$$= f_5(z(t-5) - \bar{z})$$
$$\frac{\partial z(t)}{\partial u(t-3)} = (k_7\bar{z}(t-1) + k_{11}\bar{z}(t-5))(u(t-3) - \bar{u})$$
$$= f_6(u(t-3) - \bar{u})$$
$$\frac{\partial z(t)}{\partial u(t-4)} = (k_6 + k_8\bar{z}(t-1))(u(t-4) - \bar{u})$$
$$= f_7(u(t-4) - \bar{u})$$
$$\frac{\partial z(t)}{\partial u(t-5)} = (k_9\bar{z}(t-1))(u(t-5) - \bar{u})$$
$$= f_8(u(t-5) - \bar{u})$$

where $\bar{u}$ and $\bar{z}$ are the operating points with values

$\bar{u}$ =0.3584

$\bar{z}$ =0.5261

Subtituting the values of the k's into the f's.

$f_1 = 1.376 + (0.03702-0.004919+0.004134)\times 0.3587$

$= 1.389$

$f_2 = -0.244$

$f_3 = -0.3677$

$f_4 = 0.4154 - 0.03932\times 0.5261$

$= 0.395$

$f_5 = -0.1750 - 0.03932\times 0.5261 - 0.0216\times 0.3587$

$= -0.2034$

$f_6 = (0.03702-0.02160)\times 0.5261$

$= 0.00811$

$f_7 = 0.007367 - 0.004919\times 0.5261$

$= 0.004779$

$f_8 = 0.004134\times 0.5261$

$= 0.00217$

The equation representing the operating point is of the form

$$z_{op}(t) = k_1\bar{z}(t-1) + k_2\bar{z}(t-2) + k_3\bar{z}(t-3) + k_4\bar{z}(t-4) + k_5\bar{z}(t-5) + k_6\bar{u}(t-4) + k_7\bar{u}(t-3)\bar{z}(t-1) + k_8\bar{u}(t-4)\bar{z}(t-1) + k_9\bar{u}(t-5)\bar{z}(t-1) + k_{10}\bar{z}(t-4)\bar{z}(t-5) + k_{11}\bar{u}(t-3)\bar{z}(t-5) \quad \text{(C.9)}$$

or

$$z_{op}(t) = \bar{z}(k_1 + k_2 + k_3\bar{z} + k_4 + k_5) + \bar{u}k_6 + \bar{u}\bar{z}(k_7 + k_8 + k_9) + k_{10}\bar{z}^2$$

Substituting the values of the k's, $\bar{u}$ ' s and $\bar{z}$ ' s

$$z_{op}(t) = 0.52309$$

therefore the linearized model is of the form

$$z_l(t) = z_{op}(t) + f_1(z(t-1) - \bar{z}) + f_2(z(t-2) - \bar{z}) + f_3(z(t-3) - \bar{z}) + f_4(z(t-4) - \bar{z}) + f_5(z(t-5) - \bar{z}) + f_6(u(t-3) - \bar{u}) + f_7(u(t-4) - \bar{u}) + f_8(u(t-5) - \bar{u}) \quad \text{(C.10)}$$

Rearranging the above equation

$$z_l(t) = 0.00812 + 1.389z(t-1) - 0.244z(t-2) - 0.3677z(t-3)$$

$$+ 0.3947z(t-4) - 0.2034z(t-5) + 0.008112u(t-3)$$
$$+ 0.00477u(t-4) + 0.002174u(t-5) \qquad \text{(C.11)}$$