NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

APPLICATION OF THE BOX-JENKINS TECHNIQUE FOR SYSTEM IDENTIFICATION OF AN AIRFIN HEAT EXCHANGER PROCESS

By

Joseph D. Given

A thesis submitted in partial fulfillment of the requirement for the degree of

Master of Applied Science

in the

DEPARTMENT OF CHEMICAL ENGINEERING
UNIVERSITY OF OTTAWA

Ottawa, Canada

April 20, 1987

<u>Dr. D. D. McLean</u>
Research Director

<u>Joseph D. Given</u>
Candidate

ISBN 0-315-53777-9

ACKNOWLEDGEMENTS

While laboring intensely to finish this research and complete the thesis requirement for my degree, many appropriate quotes drifted through my consciousness during the long hours and late nights,

If it's worth doing, it's worth doing well

Finish what your start

When the going gets tough, the tough get going

However the following quote attributed to president Calvin Coolidge

> "Nothing in the world can take the place of persistence. Talent will not; nothing is more common than the unsuccessful man with talent. Genius will not; unrewarded genius is almost a proverb. Education will not-the world is full of educated derelicts. Persistence and determination are omnipotent. The slogan 'press on' has solved and always will solve the problems of the human race."

seems to summarize the ambiance I experienced while attempting to complete this work.

There are many individuals who helped make this research possible and should share in the credit: my parents for their love and nurturing; my brother and sister for their interest and support; my friends (especially Phil) for their caring and encouragement. A special acknowledgement is due to Dr. David McLean for his almost divine patience, good humour and constant encouragement. However, the following says it all:

"This work was made possible through the support, hard work, perseverance, cajolement and love of my lovely wife Erin whom I hold so dear to my heart."

In conclusion I would like very much to meet that great philosopher who coined the phrase "HARD WORK NEVER KILLED ANYONE" to review together his line of reasoning. Finally, HARD WORK IS ALL IT'S CRACKED UP TO BE!!!

ABSTRACT

APPLICATION OF THE BOX-JENKINS TECHNIQUE FOR SYSTEM IDENTIFICATION OF AN AIRFIN HEAT EXCHANGER PROCESS

Joseph D. Given

The Box-Jenkins technique is a time series analysis approach to stochastic model building. The Box-Jenkins approach yields a discrete, linear, empirical stochastic model possessing maximum simplicity consistent with representational adequacy. Furthermore, the technique may be applied either on-line or off-line, under open-loop or closed-loop operation and in the presence of significant noise. Despite these advantages, the approach has not seen general acceptance or widespread use in the industrial milieu.

The effectiveness and suitability of the Box-Jenkins approach for industrial implementation was investigated by application to an airfin heat exchanger process at the Imperial Oil Refinery in Sarnia, Ontario. The current control strategy was considered inadequate; identification (modelling) of the system and development of an appropriate control strategy were deemed necessary.

The Box-Jenkins technique was demonstrated to be a simple iterative procedure which is easily implemented. However, results of the analysis of the initial data were strongly questioned. The identification procedure was repeated with new data provided by the Refinery. The results of the second analysis were consistent with expected behaviour and the physical reality of the system. The cause for the observed discrepancy in results was determined to be inadequate data and not due to any failure of the Box-Jenkins technique.

The effects of noise level, sampling rate and control rate were also assessed with computer simulations. These studies demonstrated the suitability of the approach for industrial applications and its robustness for system identification in the presence of significant noise.

A preliminary examination of the process suggested adaptive control as a possible control strategy. Due to inadequate data, a thorough investigation and possible implementation of adaptive control were not feasible. However, a review of adaptive control strategies was performed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

k, n	discrete instances of time
t	time
x, u	system input variables
$x(t)$	continuous time representation of the system input
$x_t, x(k)$	discrete time representation of the system input
y	system output variable
$y(t)$	continuous time representation of the system output
$y_t, y(k)$	discrete time representation of the system output
y_m	model output
$T[\ ]$	linear transformation
a_i, b_i	constant coefficients
$A^*(q^{-1}), B^*(q^{-1})$	polynomials of a_i's and b_i's respectively
$B, (q^{-1})$	backshift operator
V	loss function
$\mathbf{Y}$	vector of output data
Φ	matrix of input and output data
$\mathbf{b}$	vector of coefficients

$e(k)$	residual or random variable
b	estimate of coefficient vector
e	vector of residuals
$v(k)$	correlated random variable
$G^*(q^{-1})$	transfer function (in polynomial form)
W	instrumental matrix
$C^*(q^{-1})$	polynomial of coefficients c_i's
λ	maximum likelihood constant
θ	vector
$\delta(B),\omega(B),\Omega(B)$ $\psi(B),\eta(B),\rho(B)$	Box-Jenkins polynomials or operators
$v(B)$	linear filter
N_t	system noise
a_t	white noise
r, s, p, q	order of model
d	degree of differencing
∇^d	difference operator
$y_e, x_e, a_e,$ δ,ω,ρ,η	vectors

n_t	estimate of noise series
a_t	conditional residual error series
$r_{aa}(k)$	estimate of autocorrelation of time series a_t
$r_{xa}(k)$	estimate of crosscorrelation of time series x_t and a_t
T_{95}	95% system settling time
T_0	sampling rate
σ^2	variance

1.0 INTRODUCTION

Process control of a chemical plant is concerned with the operation of its various subsystems so as to optimize or maintain specified performance criteria. In practice, the control engineer is faced with the problem of identification of the system model and its parameters. To achieve this, the effects of the input variables on the output variables of interest must be generated theoretically and/or experimentally. The quality of this knowledge determines the reliability and effectiveness of the resulting control actions.

The problem of identifying variables which affect the process or system under study and the subsequent development and determination of a characteristic system model is referred to as system identification. This thesis will be concerned with system identification.

Often when studying a particular system, the system model may be derived through the application of physical and chemical principles. Models derived using theoretical relationships are classified as mechanistic. Models may also be built based upon experimentally obtained input/output data. This type of model is referred to as an empirical model. The empirical approach is an useful tool when the system, as in most practical applications, cannot be represented or modeled from first principles due to the lack of knowledge concerning some variables or the complexity of the process itself. Although the empirical approach may be applied blindly, an engineering understanding of the process, no matter what degree, is always preferable. In practice a combination of the two generally occurs. This thesis will be concerned with empirical modeling techniques.

Models, whether mechanistic or empirical, may also be classified as either deterministic or stochastic. Any model which describes the system exactly is said to be deterministic. In practice, unknown or unmeasured variables often affect the output variables thereby making the development of a deterministic representation impossible. On the other hand, it is feasible to develop a model which will predict limits having a specified probability of containing the true value of the response of interest. This type of model is said to be stochastic. Stochastic representations take into account the random nature of process measurement whereas deterministic models assume that all variations in process measurements have been taken into consideration. Since chemical plants are generally afflicted with noise, which is non-deterministic, this thesis deals with stochastic models.

Process models may be continuous or discrete. A continuous model provides estimates of the output response to given input(s) at any time t. Models which yield estimates of system output at specific time intervals are classified as discrete. Many dynamic systems fall into the discrete time category since their system outputs may only be measured at discrete instances of time (i.e. $t_0, t_0+h, t_0+2h \ldots$). Furthermore, the increasing use of digital computers for process monitoring has given rise to an abundance of discrete or sampled-data systems. For these reasons this thesis will focus on discrete system identification techniques.

Empirical modeling, which uses input and output data, that is to say a history of measurements to fit mathematical models, provides an approximate description of the input/output characteristics of the process and is therefore by definition also stochastic. Identification techniques taking into account statistical considerations such as measurement error and considering input and output variables as random variables are a branch of a wider Estimation Theory. The first modern applications of statistical estimation theory to system identification and model estimation occurred in the 1930's in the field of econometrics. The major contributions in the physical sciences to the application of these techniques for the construction and estimation of mathematical models have been by Professor G.E.P. Box and his associates.

Many of the statistical methods available in the literature for model identification and estimation are restricted to, or assume, random samples of independent observations. Sequential observations of dynamic systems rarely demonstrate independence between samples. When dependence between observations is significant or the nature of this dependence is of vital interest itself, a separate body of statistical techniques, capable of dealing with this situation is necessary for system analysis and identification. Such a body of techniques is referred to as Time Series Analysis.

The Box-Jenkins approach is one method of time series analysis developed by G.E.P. Box and G.M. Jenkins. The approach is concerned with the building of stochastic models from discrete time series in the time domain and their subsequent application in identification, prediction and control strategies. The methodology of the technique is described in reference [1]. The Box-Jenkins approach to system identification yields a discrete, linear, empirical, stochastic model possessing maximum simplicity consistent with representational adequacy. Furthermore the method is applicable to on-line identification and adaptive control.

The Box-Jenkins technique can develop statistical models for discrete data in real time without the assumption of independent sample observations. The method does not require open loop data collection and is effective in dealing with noisy and non-stationary data. Multiple input/output systems may also be modeled adequately by this approach.

Despite these appealing features, the methodology proposed by Box and Jenkins has not seen general acceptance in the industrial milieu due primarily to the lack of published industrial applications. The applications cited in the literature are for the most part simulations and lab or pilot scale operations. The major objective of the thesis is the application and evaluation of the Box-Jenkins approach using industrial data. The effectiveness of the method in developing an adequate parametric model for subsequent system control application will be the main topic of interest of this research.

The industrial system examined was an airfin heat exchanger system in operation at the Imperial Oil Refinery in Sarnia, Ontario. A schematic diagram of the system is shown in Figure 1. A feedstream of unknown volume, composition and temperature enters the airfin heat exchanger whose output stream, of known temperature only, is fed into a drum. A second process stream of unknown, uncontrolled quantity from knock-out drums also enters the same drum. The primary control objective is to maintain the drum temperature at the desired setpoint. The current control logic for achieving this objective involves a cascade control wherein the desired setpoint is maintained through

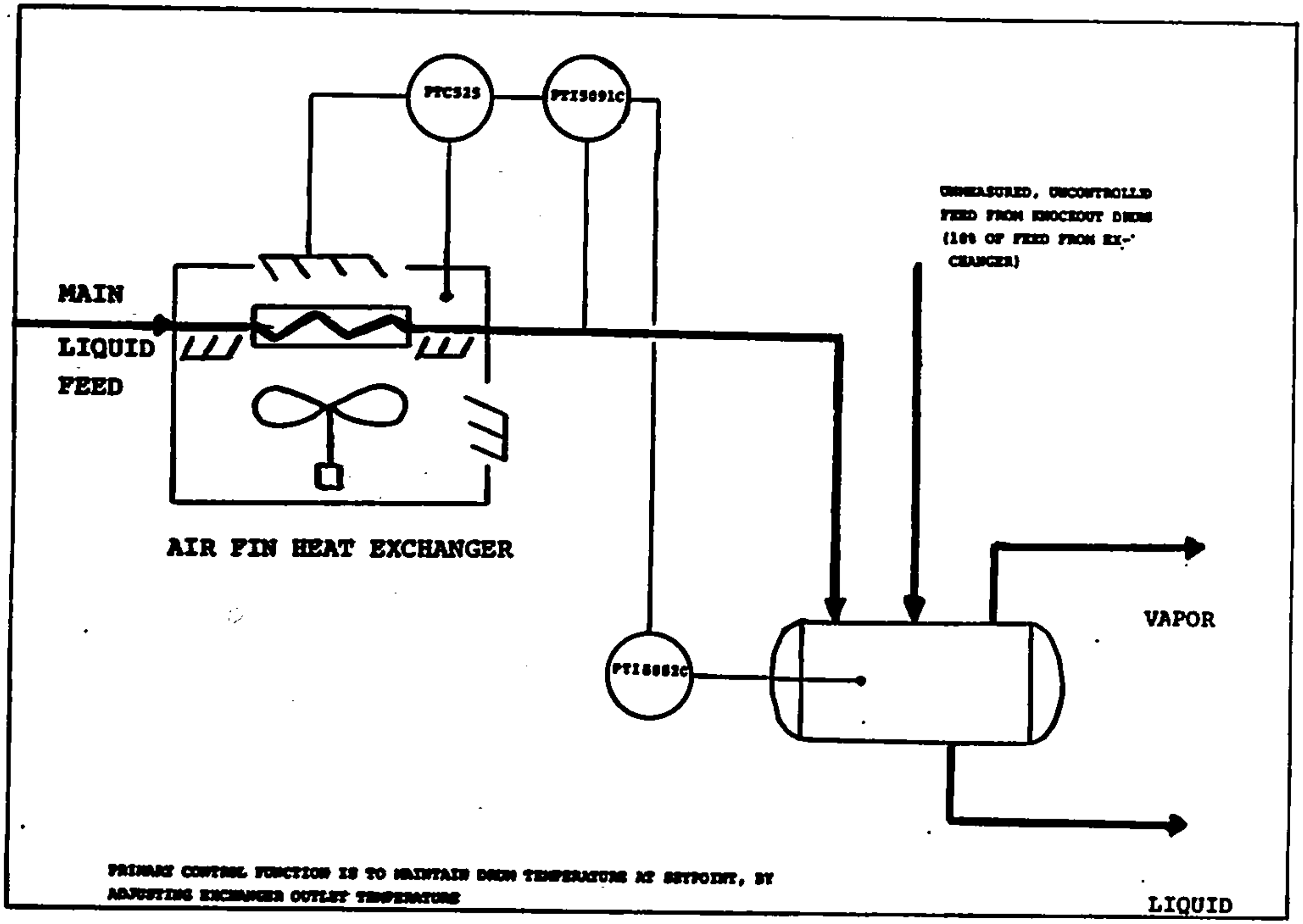


Figure 1: Schematic Diagram of the Airfin Heat Exchanger Process

manipulation of the pleneum temperature setpoint which then adjusts the position of the louvres which control air flow into the heat exchanger. Theoretically, modification of louvre position should induce a temperature change in the outlet stream leaving the heat exchanger. The control strategy is considered inadequate, since the system becomes unstable under closed loop operation and cannot maintain the drum temperature at the desired setpoint.

In summary this thesis encompasses the following objectives:

i) familiarization with and application of the Box-Jenkins approach for system identification,

ii) development of an adequate transfer function representation of the airfin heat exchanger process,

iii) simulation of the process system and determination of a control strategy which will meet the control objective,

iv) evaluation of the suitability and effectiveness of the Box-Jenkins approach for identification and control of industrial processes,

v) investigation of the feasibility of improved process control through the implementation of an adaptive control algorithm.

Initial discussions with Imperial Oil presented the possibility that adaptive control might be an appropriate alternative for control of the system. However, due to the difficulties encountered with the system, the extension to adaptive control was not undertaken but a review of the literature on adaptive control was performed and is included as Appendix 1.

2.0 SYSTEM IDENTIFICATION

The problem of system identification is generally defined as the determination of an appropriate model to describe a system or process. Eykhoff [2] defines a model as being a representation of the essential aspects of an existing system (or a system to be constructed) which presents knowledge of that system in a usable form.

It is important to emphasize the importance of the usable form of the model. For most engineers, system identification is not an end unto itself but rather a means to an end. System identification provides the necessary framework for the design of an adequate controller. Furthermore identification is closely related to prediction where the purpose of identification is to facilitate a prediction of the future behavior of the identified system.

System identification is composed of the following steps:

i) Determination of the model structure and order
ii) Estimation of the model parameters
iii) Investigation and determination of model adequacy
iv) Application of the model

Hsia [3] states that system identification can be separated into two categories: the complete identification problem and the partial identification problem. The first category deals with situations where there is a total lack of information concerning the system under study. This type of problem is also called the "black box" problem. No past knowledge or experience with the system is available, the engineer possesses no a priori information with which to guide him for the determination of model structure. Both Hsia [3] and Eykhoff [2] concur that most engineering situations do not fall into the "black box" group but rather that of the "gray box" problem or partial identification problem. The gray box problem encompasses all identification problems where the investigator possesses some a priori information of the system under consideration. Even the fact that the system is a distillation column and not a lunar vehicle implies a priori knowledge.

There exists in the literature a considerable number of system identification techniques. The purpose of this chapter is not to review all of the different methodologies available but to present the major approaches and their applicability. None of the different identification approaches that are presented can be employed to identify systems of all kinds. Therefore a classification procedure for system identification will be presented which separates the techniques based on their ability to handle different identification situations such as linearity/nonlinearity, stationarity etc. This classification will then be applied to the industrial problem under study to determine which class of available identification techniques is suitable for our problem. A detailed review of applicable methodologies will then be presented with the description and relative merits of each approach. The Box-Jenkins approach is shown through this analysis to have high potential for industrial application. It will be described in detail along with a review of published applications.

2.1 Classification of System Identification Methodologies

The following classification is partially based on reference material by Isermann [4, 5], Astrom and Eykhoff [6] and Graupe [7].

System identification methodologies have been separated according to:

i) Model type characteristics

ii) Input characteristics

iii) Identification problem characteristics

Each classification will now be presented, listing and describing briefly the individual criteria for each subset.

2.1.1 Classification by Model Type

Linear or Nonlinear Model: A system is said to be linear if the properties of proportionality and superposition are present i.e. for the following general system

$$y(t) = T[x(t)]$$

$$T[ax(t) + bx(t)] = aT[x(t)] + bT[x(t)]$$
$$= ay(t) + by(t)$$

where T denotes a transformation of x(t)

If the above equality does not hold true the system is said to be nonlinear.

Stationary or Non-Stationary Model: The distinction between stationary and non-stationary models is that the latter refers to models whose parameters vary with time. Box and Jenkins [1] state in more precise terms that a system is stationary if the joint distribution of any set of observations is unaffected by shifting all the times of observation forward or backward by an integer amount k.

Parametric or Non-Parametric Model: Systems may be characterized by a variety of representations either non-parametric or parametric. Impulse response and spectral densities are examples of non-parametric representations while differential equations are an example of parametric modeling.

The non-parametric representation of a system does not require that the order of the process be known a priori or specified explicitly. However, the non-parametric approach is intrinsically of infinite dimension. An example of a non-parametric estimation technique is spectral analysis.

Parametric techniques allow the system to be defined with a minimum number of parameters while maintaining the representational integrity of the model. But one drawback of the parametric representation is that the order of the system must be known or specified a priori. Parametric representations developed using the wrong model order may prove to be inadequate or produce results with substantial error bounds unrelated to the actual noise level of the system.

The non-structural approach for describing a system is analogous to the representation of an empirical distribution function by a histogram. They allow the system or data to "speak for themselves" and are an important first step in their respective analyses. Furthermore, the non-parametric approach may involve the estimation of many lag correlations and/or many spectral ordinates where a parametric model of one or two parameters could represent the system. Each correlation and each spectral ordinate is itself a parameter to be estimated. Therefore, non-parametric approaches might be very prodigal with parameters whereas the parametric approach could be parsimonious. Some parametric methods incorporate a non-parametric procedure as a first step or tool for the identification of the model order thereby gaining the best of both worlds, the correct model order and a parsimonious parameter representation of the system. The Box-Jenkins technique is an example of this approach.

Deterministic or Stochastic System Model: If the system can be predicted exactly, it is said to be a deterministic representation of the system. However, in most practical situations, the system under study requires a stochastic model, in that the future is only partly determined by the measured variables. Deterministic models assume that all variation in process measurements have been accounted for, while stochastic models take into consideration the random nature of process measurement.

Continuous or Discrete Model: A system may be either discrete or continuous. When the input and output variables are known for any instant of time, the system is said to be continuous. The discrete case occurs when the system is sampled at specific instances of time (usually equispaced intervals). The transformation from the continuous model representation to a discrete formulation is usually straightforward [7].

2.1.2 Classification by Input Class

Open Loop or Closed Loop Identification: Identification of the dynamic parameters of a system is valid only if the measurements or data reflect a transient state in the system. Dynamic parameters cannot be identified by any technique when the system is at steady state.

Since, during normal operation in industry, a system is usually subject to control to maintain a certain level of operation, variations in both the input and output variables are small. Therefore normal operating data quite often does not sufficiently excite the system to allow good identification of the dynamic parameters. This situation then requires the use of induced perturbations either deterministic such as a step, impulse or sine wave input, or random such as a white noise input.

The earliest or classical methods of engineering identification are those based on the use of deterministic inputs and measurement of the system response. Since special inputs rather then normal operation inputs are necessary, it is obvious that the above techniques must be performed under open loop conditions. Some authors, such as Graupe [7], refer to this condition as off-line identification but the use of the term off-line identification in this text will be reserved for batch processing identification (discussed in a subsequent section). In the industrial milieu, in general, no system is operating independently or uncontrolled therefore the condition of open loop implies non-normal operating conditions or the off-line situation as described by Graupe [7]. The advantage of the special input signals is that significant simplifications in the computations can be achieved.

These techniques however have certain disadvantages which make them unattractive and difficult to implement for the industrial case. Firstly, the condition of open loop operation may render the system unstable which is unacceptable. Secondly, the perturbations used must be of a magnitude that is both relevant and tolerable by the system. These restrictions on input variation may not produce system responses large enough to be distinguished from the inherent noise in the system. Therefore the deterministic input approach assumes a noise-free system or the use of high level inputs necessary to override the noise, both of which may be unacceptable alternatives. Finally this approach assumes linear stationary processes such that the input/output relations are consistent for one or all sets of input(s). Linearity and stationarity of the system may and generally are assumed over the normal operating range of the system but cannot be extended in most cases over the total possible range of operation of the system thereby implying that models derived under open loop conditions may not adequately describe the system under normal closed loop operation.

The use of random input functions and the correlation function approach facilitates system identification under both open and closed loop operation. These techniques have several characteristics which are appealing for industrial application. One advantage is the option of using open or closed loop operating conditions during data collection. The use of closed loop identification alleviates several of the problems associated with the open loop condition such as stability of the system, and the validity of the model under restricted linearity conditions since the data collection is performed such that the model will be valid over the normal operating ranges of the system.

Furthermore the assumption of a noise free system is not required and the level or variance of the white noise input need not be exceedingly high to obtain good identification. The one drawback for closed loop identification using this approach is that the control scheme be adequate and the system be stable during closed loop operation. However, if these conditions are not present, the advantages of using closed loop data over open loop data become almost insignificant since the stability of the system is not guaranteed in either case. Closed loop identification is discussed in more detail by MacGregor [8] and Astrom and Eykhoff [6].

Off-line or On-line Identification: As a result of the use of process computers a system identification scheme can be performed either off-line or on-line.

For off-line identification the data collection and identification analysis are performed separately. Measurements are taken during real time, stored and transferred for evaluation. Batch processing, i.e. all the data are analyzed at once, is usually applied which allows for direct, or one shot, identification schemes.

On-line or real time identification involves using process data sequentially as it becomes available. The analysis may be performed after each sample of data is measured or on sequential finite sets of data.

Single or Multi-input System: There is a straightforward and obvious distinction between the single and multi-input situations. This classification is of interest due to the fact that identification techniques are considerably simplified if the system is affected by only one input, rather then a combination of several simultaneous disturbances or inputs. Also, several identification techniques, such as the majority of the classical techniques, cannot handle multi-input system identification.

2.1.3 Identification Problem Characteristics

This classification may be the most important and least definable of the three presented, the classification category being the amount or degree of a priori information concerning the system to be identified. It also implies the degree of difficulty associated with a given identification problem. All of the previously discussed classifications depend upon the available a priori knowledge of the system. The a priori knowledge or lack thereof directly influences the data collection and choice of identification technique. Identification approaches that presume or require less a priori knowledge are, of necessity, less accurate and are more mathematically involved and complex while needing more computational power then those techniques which assume more prior knowledge of the system.

2.1.4 Classification of the Airfin Heat Exchanger Identification Problem

The data provided by Imperial Oil are discrete and gathered during open loop-operation. The input signal is a random time series. Although the initial data are open loop, the final goal of the investigation is the implementation of an adequate, stable controller which may necessitate closed-loop identification at a later stage for verification of model adequacy during normal operating conditions. The available a priori knowledge is very limited. Thus a mechanistic model is out of the question and an empirical representation is required.

It is assumed that a linear model representation will accurately describe the system over the desired operating range although the system may possess some non-linear characteristics. This assumption is common practice [9].

Due to unmeasured disturbances in both flow and temperature the condition of stationarity is not a valid assumption since system parameters are expected to be time varying to some degree. A stochastic representation is necessitated due to the unknown nature of several key variables which affect the output. The initial investigation is being performed off-line although on-line identification will be required for closed-loop identification and the implementation of the adaptive control scheme should it prove feasible.

Although the current system is a single input-single output configuration, this investigation may provide just cause for the consideration of additional input variables to provide adequate control.

A parametric representation is deemed appropriate, to facilitate implementation of a control law. The presence of noise in the system is significant and of an unknown nature, that is to say it cannot be assumed to be white noise.

To summarize, the identification procedure required for this study must be able to develop or build a linear, parametric, discrete stochastic model with limited a priori knowledge. Furthermore, the method must be capable of performing this analysis under open-loop or closed-loop operation, either off-line or on-line in the presence of significant noise and possible nonstationarities in the system. An additional requirement of the approach is the possibility of extension to a multi-input configuration.

2.2 Review of Identification Techniques

2.2.1 Introduction

Of the wide assortment of available techniques for system identification only those techniques which are applicable to the airfin heat exchanger problem are covered here. The subsequent review of these approaches will demonstrate the motivation for using the Box-Jenkins approach.

2.2.2 Review of Applicable Identification Techniques

The criteria for selecting an appropriate technique were outlined in the previous section, the review of the literature for appropriate techniques is the next step.

The classical methods of system identification are inappropriate for application here. Step response identification has the step function as its special input. But its restrictions include the need for adequate filtering of noise and the assumed stationarity of the system over the response interval exhibited by the system. The frequency response technique is based on the work by Nyquist [10] and Bode [11]. But this technique is only suitable for linear, stationary, open-loop identification and the presence of noise greatly undermines its accuracy. The impulse response approach suffers from the same inadequacies, further heightened by its inability to handle identification of linearized forms of nonlinear systems [7]. None of the classical methods can cope well with noise since each approach ignores the presence of noise by assuming a deterministic noise-free system [8].

The following identification techniques drawn from the literature were considered acceptable alternatives for the problem under investigation using the selection criteria of section 2.1.4:

i)	LEAST SQUARES REGRESSION	(LS)
ii)	GENERALIZED LEAST SQUARES	(GLS)
iii)	INSTRUMENTAL VARIABLES	(IVA)
iv)	MAXIMUM LIKELIHOOD	(ML)

v) STOCHASTIC APPROXIMATION (STA)

vi) CORRELATION AND LEAST SQUARES (COR)

The identification techniques will be presented and outlined individually with a short description and evaluation of their respective advantages, disadvantages and suitability vis a vis the identification problem. This analysis will identify the Box-Jenkins approach as the optimal choice of available system identification techniques.

Least Squares Regression (LS)

Consider a linear, time-invariant, discrete-time system with a single input and single output. It is assumed that the system can be represented by the following model notation called the canonical representation.

$$y(k) + a_1 y(k-1) + \ldots + a_n y(k-n) = b_1 u(k-1) \ldots + b_n u(k-n) \quad (1)$$

In equation (1) $u(k)$ is the value of the input and $y(k)$ is the value of the output. The a's and b's are the model coeffecients.

Given the following definitions:

$$A^*(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2} + \ldots + a_n q^{-n} \quad (2)$$

$$B^*(q^{-1}) = b_1 q^{-1} + b_2 q^{-2} + \ldots + b_n q^{-n} \quad (3)$$

equation (1) may be written as

$$A^*(q^{-1}) y(k) = B^*(q^{-1}) u(k) \quad (4)$$

where is identical to the backshift operator B as defined by Box and Jenkins [1].

The loss function is a quantitative measure of the equivalence of the model output given a particular set of parameters with the measured system output. The loss function V is a function of the process output y, and the model output, y_m .

$$V = V(y, y_m)$$

The goodness of fit is therefore optimized by making the loss function as small as possible.

The loss function for the Least Squares approach is

$$V = \sum_{k=n}^{N+n} e^2(k) \tag{5}$$

where $e(k)$ is the generalized error or residual defined by

$$e(k) = (y(k) - y_m(k)) \tag{6}$$

To determine the minimum of the loss function we introduce

$$Y = \begin{bmatrix} y(n+1) \\ y(n+2) \\ \cdot \\ \cdot \\ y(n+N) \end{bmatrix} \qquad \Phi = \begin{bmatrix} -y(n) & -y(n-1) & \dots & -y(1) & u(n) & \dots & u(1) \\ -y(n+1) & & \dots & -y(2) & u(n+1) & \dots & u(2) \\ & & \cdot & & & \cdot & \\ & & \cdot & & & \cdot & \\ & & \cdot & & & \cdot & \\ -y(n+N-1) & & \dots & -y(N) & u(n+N-1) & \dots & u(N) \end{bmatrix}$$

$$b^T = [a_1 \, a_2 \dots a_n \, b_1 \, b_2 \dots b_n]$$

Then the error equation (6) reduces to

$$e = y - \Phi \hat{b} \qquad (7)$$

If $\Phi^T \Phi$ is not singular then the loss function is minimized when

$$\hat{b} = (\Phi^T \Phi)^{-1} \Phi^T y \qquad (8)$$

A more detailed discussion of Least Squares estimation can be found in Hsia [3]. Press et al [7] state that Least Squares fitting is a maximum likelihood estimation of the fitted parameters if the measurement errors are independant and normally distributed with constant standard deviation. Isermann [5] states that the Least Squares technique can be used as a recursive on-line parameter estimation tool. Hsia [3] discusses the application of the real time Least Squares algorithm for tracking time varying parameters. One requirement of this approach is that the residuals must be uncorrelated to obtain unbiased parameter estimates. Correlated residuals can provoke erroneous results therefore the method is critically dependent upon the assumption of uncorrelated residuals. Hsia [3] and MacGregor [8] conclude that the Least Squares approach is suitable for systems with small noise disturbances only, where the condition of white noise can be assumed. This condition fails in the presence of strong or coloured noise disturbances. Another serious drawback is that the a priori choice of model order is required to prevent erroneous results. Astrom and Eykhoff [6] state the initial choice of model order is critical. Finally the Least Squares approach assumes stationarity or quasi-stationarity of the system parameters over the measurement period.

Generalized Least Squares (GLS)

The Generalized Least Squares (GLS) technique is a modification of the (LS) approach developed to eliminate the necessary condition of uncorrelated residuals. The (GLS) methodology can provide more accurate estimates in the presence of correlated residuals.

Given the following system

$$A^*(q^{-1})\,y(k) = B^*(q^{-1})\,u(k) + v(k) \tag{9}$$

$y(k)$ and $u(k)$ are the system output and input respectively and $A^*(q^{-1})$ and $B^*(q^{-1})$ are as previously defined in equations (2) and (3). $v(k)$ is a sequence of correlated random variables. We assume that $v(k)$ can be represented as

$$v(k) = G^*(q^{-1})\,e(k) \tag{10}$$

Where $e(k)$ is an uncorrelated random variable or residual and $G^*(q^{-1})$ is the transfer function. The individual weights in $G^*(q^{-1})$ are called the impulse response function.

Substituting equation (10) into equation (9) yields the following equation

$$A^*(q^{-1})\,y(k) = B^*(q^{-1})\,u(k) + G^*(q^{-1})\,e(k) \tag{11}$$

Define the following variables

$$\tilde{y}(k) = \frac{1}{G^*(q-1)} y(k) \qquad (12)$$

$$\tilde{u}(k) = \frac{1}{G^*(q-1)} u(k) \qquad (13)$$

Substituting equations (12) and (13) into equation (11) the equation becomes

$$A^*(q^{-1})\, \tilde{y}(k) = B^*(q^{-1})\, \tilde{u}(k) + e(k) \qquad (14)$$

Therefore, if $\tilde{y}(k)$ and $\tilde{u}(k)$ are now considered the system output and input, the identification problem is reduced to an ordinary Least Squares problem. This is evident when one compares equation (14) with equation (6). The (GLS) approach can be interpreted as a (LS) identification problem with the same loss function, equation (5), but with the generalized error defined as

$$e(k) = \frac{1}{G^*(q-1)} (y(k) - y_m(k)) \qquad (15)$$

In practice, correlation of the residuals and the transfer function $G^*(q^{-1})$ are seldom known a priori. Clarke [12] has proposed an iterative procedure for determining the impulse transfer function. This iterative approach essentially produces a (LS) fit with uncorrelated residuals. The (GLS) approach overcomes a major drawback of the (LS) technique but increases the degree of a priori knowledge required. A further disadvantage stated by Astrom and Eykhoff [6] is the absence of any systematic criteria for the choice of order for both the system and noise models. Another major deterrent is the lack of a theoretical proof of convergence for Clarke's technique although it has been successfully implemented with an ad hoc choice of model order in specific examples. Unreliable convergence of the (GLS) technique is the conclusion drawn by Isermann et al [13] in their review of six on-line identification schemes. Isermann [5] also reports the possibility of biased estimates using (GLS).

Instrumental Variables (IVA)

When the system dynamics are the only focus of the analysis and a model of the environmental effects, i.e. disturbances of various kinds other than the identified system input which corrupt the system output, is not deemed a necessity, there exists a technique for determining the system model while avoiding the difficulties associated with correlated residuals. The approach is called Instrumental Variables (IVA).

The (LS) estimation

$$\hat{b} = (\Phi^T \Phi)^{-1} \Phi^T y \qquad (16)$$

is obtained by premultiplying equation (7) with Φ^T and neglecting $\Phi^T e$. The estimate $\hat{b}$ will be unbiased if the term $\Phi^T e$ has a zero mean. The mean is non zero when the residuals are correlated.

The (IVA) approach modifies equation (7) by multiplication with W^T , where W is the instrumental matrix whose elements are functions of the measured data with the following properties:

$$E(W^T \Phi) \quad \text{nonsingular} \tag{17}$$

$$E(W^T e) \quad \text{is zero} \tag{18}$$

where E is the expected value

The parameter estimates obtained from the (IVA) equation will then be :

$$W^T y = W^T \Phi \hat{b} + W^T e \tag{19}$$

Substitution of equation (18) into the above equation yields the following equation for the parameter estimates.

$$W^T y = W^T \Phi \hat{b} \tag{20}$$

The parameter estimates obtained from equation (20) will then be unbiased.

The (IVA) approach, like the GLS technique, is a modified derivation of the original LS technique. This methodology was developed to perform system identification in the presence of substantial noise. The (IVA) approach will produce unbiased estimates of the system dynamics, when properly implemented, while ignoring system disturbances. No information concerning the noise characteristics can be derived from this method. A further disadvantage reported by Isermann et al [13] is the poor reliability exhibited by the technique. He further states that the degree of a priori knowledge for implementation is significantly increased by the necessity of choice of the initial instrumental matrix W. Isermann [5] further concludes, in a later article, that the (IVA) approach possesses no reliable convergence.

Maximum Likelihood Estimation (ML)

The maximum likelihood approach postulates that the system under study possesses correlated residuals. The single input-single output system can be represented by

$$A^*(q^{-1})\, y(k) = B^*(q^{-1})\, u(k) + \lambda C^*(q^{-1})\, e(k) \qquad (21)$$

where $A^*(q^{-1})$, $B^*(q^{-1})$, $y(k)$, $u(k)$ and $e(k)$ are as previously defined. $C^*(q^{-1})$ is of the same form as $A^*(q^{-1})$ and $B^*(q^{-1})$ (i.e. a polynomial).

The parameters in equation (21) can be estimated using maximum likelihood theory. The likelihood function L is given by

$$-\log L(\theta,\lambda) = (1/2\lambda^2) \sum_{k=1}^{N} \epsilon^2(k) + (N/2)\log\lambda + (N/2)\log 2\pi \qquad (22)$$

$$C^*(q^{-1})\, \epsilon(k) = A^*(q^{-1})\, y(k) - B^*(q^{-1})\, u(k) \qquad (23)$$

$$\epsilon(k) = \lambda\, e(k) \qquad (24)$$

The likelihood function is considered as a function of θ , a vector whose components are the parameters $a_1 a_2 \ldots a_n, b_1 b_2 \ldots b_n$, $c_1 c_2 \ldots c_n$, the initial conditions of equation (23) and λ . The log of the likelihood function is linear in the parameters a_i and b_i but strongly nonlinear in c_i . The optimization of L with respect to θ and λ can be performed separately. First θ is determined such that the loss function

$$V(\theta) = \sum_{k=1}^{N} \epsilon^2(k) \qquad (25)$$

is minimized with respect to θ . The optimization of L

with respect to λ is then performed analytically yielding the following equation

$$\lambda^2 = (1/N) \min_\theta V(\theta) \tag{26}$$

The (ML) estimates possess desirable asymptotic properties and estimates of parameter precision can also be calculated (refer to Astrom, Bohlin and Wensmark [14] for a detailed discussion). The (ML) approach may be considered as a (GLS) method whose filter function, G = 1/C , is determined automatically [6]. Bohlin [15] states that the estimates are consistent, asymptotically normal and efficient (i.e. minimum variance) under the conditions that the likelihood function is correct, the input and process disturbances are uncorrelated and that the system is "persistently excited". Saridis [16] states that, for stable systems, a proof of convergence for the on-line algorithm is possible but initial parameter estimates must be close to the true values to guarantee convergence. He also adds that the assumption of a Gaussian distribution for all random variables involved is a major disadvantage of the method. Bohlin [15] supports these arguments and further adds that the prerequisites for successful identification are not always known to be satisfied a priori. Isermann [5] reports a slow convergence rate for the (ML) recursive procedure.

Stochastic Approximation (STA)

Stochastic Approximation techniques (STA) were developed to provide system identification while taking into consideration the stochastic nature of the sampled data, that is to say, the presence of measurement error. The available recursive (STA) algorithms are generally based on the Robbins-Monro or Kiefer-Wolfowitz algorithms [3]. These basic (STA) techniques yield biased estimates if the generalized error is used for minimization. Saridis and Stein [17] demonstrated that unbiased estimates of the impulse response are possible by minimizing the output error. The algorithm proposed by Saradis and Stein converges in the mean square sense to the true parameter values if the following conditions hold:

i) the input is an independent random variable with expected value of the mean equal to zero

ii) the noise is an independent random variable with expected mean value equal to zero

Isermann et al [13] demonstrated a very good performance with large identification time periods for the (STA) approach however the overall reliability of the technique is poor and the necessary a priori knowledge is considerable. They also reported that short identification time periods produce results with large variances.

Correlation Analysis with Least Squares Parameter Estimation (COR)

A major drawback of the (LS), (GLS), (ML) and (IVA) techniques is the required a priori knowledge or assumption of model structure and order before implementation. The (COR) approach eliminates this requirement by utilizing a two step identification procedure which first identifies model order using a nonparametric stage, then performs a parameter estimation procedure on the identified model.

Autocorrelation and crosscorrelation function analysis provide the methodology used to determine or estimate model order. Parameter estimates are then calculated using Least Squares. A more detailed discussion of parameter estimation is found in Isermann et al [13]. Definitions of the autocorrelation and crosscorrelation functions, will be presented in a later chapter.

Correlation analysis with Least Squares estimation is a two-step system identification technique which can be implemented both off-line and on-line. The Box-Jenkins approach using time series is a similar two stage scheme of identification and estimation although the maximum likelihood approach rather than Least Squares is used for parameter estimation. Isermann et al [13] list several advantages of the (COR) approach: very little a priori knowledge about the system and model structure is required; the method requires small storage capacity and minimum computational expense; structure, order and time delay of the system model are easily determined; intermediate results are accessible for evaluation of performance; initial estimates of matrices and parameters are unnecessary; accuracy of results and computational expense can be easily manipulated to produce either greater accuracy with increased computations or reduced computational expenditure at the cost of accuracy of the results; and finally divergence is not possible.

Astrom and Eykhoff [6] also report that the a priori knowledge required for implementation is minimal. Isermann [5] further states in a later paper that the (COR) approach exhibits good performance for a wide class of noise models and possesses a simple procedure for verification of model adequacy. Isermann et al [13] conclude that for general linear processes (COR) presents the most advantages in comparison with six other identification techniques.

In addition to these favorable characteristics, the Box-Jenkins approach also possesses the following advantages: the assumption of uncorrelated observations is not mandatory for successful implementation; the model identified possesses maximum simplicity consonant with representational adequacy; the technique does not require open loop operation, if a known, artificially generated noise signal is added, and is effective in the presence of considerable noise; the condition of stationarity of the system is not required. Furthermore, the Box-Jenkins approach can be extended to multi-input configurations. Chatfield [18] states the following with respect to the Box-Jenkins approach:

> "...these authors show how to identify a linear parametric model for a system and hence find a "good" control procedure. It is interesting to note that a somewhat similar approach has been developed by two control engineers in Astrom and Bohlin, though the latter do not give anywhere near such detailed guidance on estimation, diagnostic checking and the treatment of non-stationary series."

The method developed by Astrom and Bohlin is described in reference [19].

Summary

In summary the Box-Jenkins approach is considered the most appropriate technique for resolving the identification problem outlined in section 2.1.4. The (LS) technique is inappropriate since it assumes stationarity of the data and disturbances that are uncorrelated (i.e. white noise). The (GLS) approach removes the assumption of uncorrelated residuals but requires a greater degree of a priori information. Due to the possibility of biased estimates, the non-existence of a systematic approach for determining model order of both the system and noise models and unreliable convergence the (GLS) is rejected as a plausible alternative. The (IVA) technique which provides no information concerning system noise characteristics, is removed from consideration due to its poor reliability, inability to handle nonstationarities and the considerable amount of a priori knowledge required. The (ML) approach requires initial estimates of the parameters which are close to the true values. This necessary condition and the uncertainty of whether or not all the prerequisites of the technique are satisfied before implementation justify its elimination from consideration. The (STA) approach is not feasible for implementation due to its poor reliability and the considerable extent of required a priori knowledge of the system.

The Box-Jenkins approach fulfills all the necessary requirements listed in section 2.1.4. The method requires little a priori knowledge and is effective in the presence of considerable noise. The method allows identification of model order and encourages parsimony in parameterization of the model. The assumption of stationarity is not required for effective implementation. The procedure may be performed off-line or on-line using open or closed loop data collection. Furthermore, it can be extended to handle multi-input configurations. Therefore the Box-Jenkins approach will be used to perform the system identification of the problem under study.

2.3 The Box-Jenkins Identification Procedure

The Box-Jenkins technique is concerned with the development of stochastic models in the time domain for discrete time series. Possible areas of application include identification, description, prediction and control. The area of interest for this study is the development of an optimal control scheme for the airfin heat exchanger system.

2.3.1 The Box-Jenkins Transfer Function-Noise Model Representation

Box and Jenkins represent a discrete dynamic system, that is to say, the dynamic relationship between input X and output Y measured at equispaced instants in time, by a difference equation. The system output Y_t at time t is modelled as the linear combination of past values of the input and output. The linear model may be written in the following form using the backshift operator B which is defined by

$$B X_t = X_{t-1}$$

$$B^m X_t = X_{t-m}$$

where X_t is a discrete observation at time t of a time series X :

$$(1-\delta_1 B-\ldots-\delta_r B^r)\, Y_t = (\omega_0-\omega_1 B-\ldots-\omega_s B^s)\, X_{t-b} \quad (27)$$

If we define the following two operators $\delta(B)$ and $\omega(B)$

$$\delta(B) = (1-\delta_1 B-\ldots-\delta_r B^r) \quad (28)$$

$$\omega(B) = (\omega_0-\omega_1 B-\ldots-\omega_s B^s) \quad (29)$$

we can then write equation (27) in the following form

$$\delta(B)\, Y_t = \omega(B)\, B^b X_t \quad (30)$$

$$\delta(B)\, Y_t = \Omega(B)\, X_t \quad (31)$$

If Y_t and X_t represent deviations at time t from appropriate mean values, or equilibrium, then to an adequate approximation, the system can be represented by a linear filter of the form

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots \qquad (32)$$

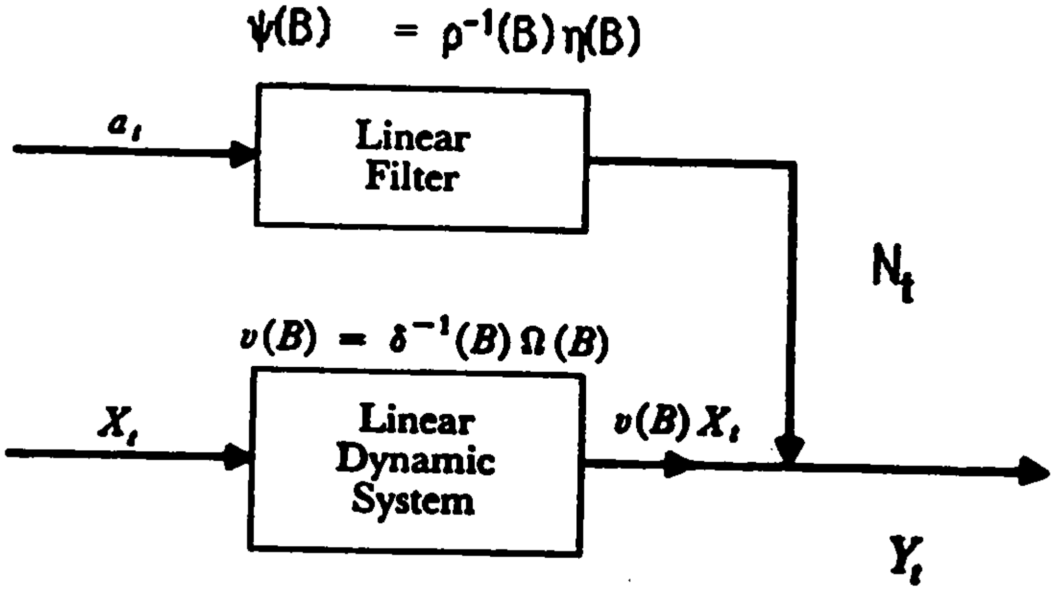
$$Y_t = (v_0 + v_1 B + v_2 B^2 + \dots) X_t$$

$$Y_t = v(B) X_t$$

The output deviation at some time t is represented as a linear aggregrate of input deviations at time $t, t-1, t-2, \dots$. The operator $v(B)$ is defined as the transfer function of the filter and hence of the dynamic system itself. The weights $v_0, v_1, v_2, v_3, \dots$ in equation (32) are called the impulse response function of the system. Furthermore, the transfer function $v(B)$ can be expressed as the ratio of the polynomials $\Omega(B)$ and $\delta(B)$.

$$v(B) = \Omega(B) / \delta(B) = \delta^{-1}(B)\Omega(B) \qquad (33)$$

The Box-Jenkins approach for estimating an appropriate model, relating an output Y and an input X , is identical to estimating the transfer function $v(B) = \delta^{-1}(B)\Omega(B)$. However the problem is compounded in practice since the output Y_t cannot be expected to follow exactly the pattern determined by the transfer function even if the model used is appropriate and adequate. Disturbances of various kinds other than X_t ordinarily corrupt the system. Box and Jenkins consider disturbances, which can occur anywhere in the system, in terms of the net effect on the output Y_t , as shown in Figure 2. They further assume the net efffect, or noise N_t , is independant of X_t and is



$$Y_t = v(B)X_t + N_t$$

Figure 2: The Box-Jenkins transfer Function-Noise Model

additive with respect to the influence of X_t on the output Y_t .

$$Y_t = \delta^{-1}(B)\Omega(B)X_t + N_t \tag{34}$$

The noise N_t can be described by the following linear filter with the input time series being a white noise process.

$$N_t = \psi(B)a_t = \rho^{-1}(B)\eta(B)a_t \tag{35}$$

The white noise process will be defined and discussed in more detail in a later section.

It is necessary therefore, in practice, to estimate the transfer function $\psi(B)$ of the linear filter describing the noise and the transfer function $v(B)$ describing the dynamic relationship relating the input and output variables.

However, both the transfer functions $v(B)$ and $\psi(B)$ can adequately be represented by a ratio of polynomials of low degree in B that is to say containing only a few parameters ($r \leq 2$, $s \leq 2$ for the dynamic model and $p \leq 2$, $q \leq 2$ for the noise model). Using the model form of a ratio of polynomials allows the identification and estimation of the smallest number of parameters necessary for adequate representation. The principle of parsimony in the use of parameters is extremely important since overparameterization of the model form can lead to unnecessarily poor estimation of the output Y_t .

A further complication encountered when trying to model a dynamic system in practice is the presence of nonstationary behaviour. Many time series in industry exhibit this type of behaviour, and, in particular, do not vary about a fixed mean. But although the fluctuations occur about different levels for different times, the broad behaviour, taking into account changes in level, may be similar. Box and Jenkins represent such a system by a model which calls for the dth difference of the process to be stationary. Using the noise model as an example.

$$\rho(B)\, N_t = \phi(B)(1-B)^d N_t = \eta(B)\, a_t \tag{36}$$

that is

$$\phi(B)\, n_t = \eta(B)\, a_t$$

where

$$n_t = (1-B)^d N_t = \nabla^d N_t$$

In practice d is usually 0, 1, or at most 2.

The model form of equation (36) describing the relationship between a time series, N_t in this case, and a white noise process is called an autoregressive integrated moving average (ARIMA) model. Equation (36) provides a powerful model for describing stationary and non-stationary times series. The order (p,d,q) of the ARIMA model is defined by the number of autoregressive terms p, the degree of difference required to induce stationary behaviour d, and the number of moving average terms q

$$(1-\phi_1 B - \ldots - \phi_p B^p)(1-B)^d N_t = (\eta_0 - \eta_1 B - \ldots - \eta_q B^q)\, a_t \tag{37}$$

In summary, Box and Jenkins represent a dynamic system contaminated with some degree of noise by the following equality

$$Y_t = v(B)X_t + N_t \tag{38}$$

where Y_t , $v(B)$, X_t and N_t are as previously defined. Successful adequate representation of the system requires the identification and estimation of both the dynamic transfer function model and the noise model.

2.3.2 Building the Transfer Function-Noise Model

The Box Jenkins approach to model building is an interactive one comprising preliminary identification, model fitting or parameter estimation and diagnostic checks. The preliminary identification procedure is designed to suggest an appropriate model form or order. The estimation stage provides maximum likelihood estimates of the system parameters. Quite often the initial model form will prove to be an inadequate representation of the system and hence the third step of the model building approach, diagnostic checking, has been developed to not only detect model inadequacy but to also suggest appropriate modifications for a further iterative cycle. A short description of each stage of this iterative procedure is now presented to provide some additional insight into the technique. The following sections are by no means intended to be an in-depth analysis of the Box-Jenkins approach and interested readers are referred to chapter 11 of Box and Jenkins [1] for a detailed discussion of the procedure and relevant theory.

Preliminary Identification

The preliminary identification of the combined transfer function-noise model involves the following steps:

i) Calculation of the autocorrelation function of the input time series,

ii) Determination of the pre-whitening operator for the input time series,

iii) Application of the pre-whitening operator to both the input and output time series,

iv) Calculation of the cross correlation function of the transformed input and output time series from step iii),

v) Preliminary identification of tentative model orders, r, s and the delay parameter b,

vi) Calculation of initial estimates of the preliminary model form,

vii) Computation of an estimate of the noise time series using the preliminary transfer function model,

viii) Calculation of the autocorrelation function of the estimated noise series,

ix) Identification of preliminary tentative model orders, p, q and degree of differencing d required to induce stationarity of the estimated noise series,

x) Computation of initial parameter estimates of the preliminary noise model.

In summary, the preliminary identification stage provides a tentative class of models to be investigated and the approach commits the user to nothing more. Intuition or a "gut feeling" concerning model type may be superimposed, substituted or used in conjunction with the preliminary identification outlined in this section.

A description of the autocorrelation and crosscorrelation functions and sample calculations may be found in Appendix 2. Prewhitening of the input is discussed in Appendix 3 but its primary advantage is that when implemented, considerable simplification of the identification procedure is achieved.

Estimation

Having identified a tentative formulation for the model, efficient estimates of the parameters are now needed. R.A. Fisher [1], page 208, stated that efficient use of the data for parameter estimation is required to permit relevancy of diagnostic checks. If this requirement is not present inadequacy of model fit may be due to inefficient fitting and not incorrect model form selection.

The combined transfer-function-noise model must be estimated simultaneously even if the system noise model is not of interest. However, for stochastic control both models weigh equally in importance.

If the starting values y_0, x_0, a_0, where x_0, y_0 and a_0 are vectors with values of the input, output and white noise series prior to the beginning of the series are known then given the data with any choice of parameters

$$(b,\delta_1,\ldots,\delta_r,\omega_0,\omega_1,\ldots,\omega_s,\rho_1,\rho_2,\ldots,\rho_p,d,\eta_1,\eta_2,\ldots,\eta_q),$$

it is possible to calculate successive values of a_t where $a_t = a_t(b,\delta,\omega,\rho,\eta \mid x_0,y_0,a_0)$ for $t = 1,2,\ldots n$. Assuming the calculated series a_t to be normally distributed white noise then a close approximation to the maximum likelihood estimates of the true parameters [1] may be obtained by minimization of the conditional sum of squares of residuals.

$$S_0(b,\delta,\omega,\rho,\eta) = \sum_{t=1}^{n} a_t^2(b,\delta,\omega,\rho,\eta \mid x_0,y_0,a_0) \qquad (39)$$

where

$$\delta = (\delta_1,\ldots,\delta_r)$$
$$\omega = (\omega_0,\omega_1,\ldots,\omega_s)$$
$$\rho = (d,\rho_1,\rho_2,\ldots,\rho_p)$$
$$\eta = (\eta_1,\eta_2,\ldots,\eta_q)$$

Minimization of equation (39) yields conditional maximum likelihood estimates. In practice x_0, y_0 and a_0 are unknown and initial values of the input and output time series are substituted for x_0 and y_0 respectively and prior values of a_0 assume their expected value of zero.

The calculation of the estimated residual time series involves the following steps:

i) Estimation of the output time series using equation (40)

$$\hat{y}_t = \delta^{-1}(B)\,\omega(B)\,X_{t-b} \tag{40}$$

ii) Calculation of the noise series estimate $\hat{n}_t$ using equation (41)

$$\hat{n}_t = y_t - \hat{y}_t = y_t - \delta^{-1}(B)\,\omega(B)\,X_{t-b} \tag{41}$$

iii) Generation of the conditional residual error time series using equation (42)

$$\hat{a}_t = \eta^{-1}(B)\,\rho(B)\,\hat{n}_t \tag{42}$$

The estimated residual error time series calculated for a conditional set of parameters is then used for diagnostic checking of model adequacy.

<u>Diagnostic Checking of Model Adequacy</u>

Identification and estimation of the combined transfer function-noise model having been accomplished the application of diagnostic checks is now appropriate. Specific checks are implemented using the residuals of the fitted model which allow the data itself to suggest modifications to the model. The procedure of overfitting that is to extend the model in a particular direction is also applicable here but assumes some a priori knowledge of expected discrepancies.

Serious model inadequacy of the combined transfer function-noise model can usually be detected by analysis of the residuals.

The autocorrelation function $r_{aa}(k)$ of the estimated residual error time series should behave as that for white noise, that is to say, no marked correlation patterns should be present. The crosscorrelation function $r_{xa}(k)$ of the input time series and estimated residual error time series should display no significant non zero terms. If these two conditions hold true then the combined transfer function/noise model is considered to be an adequate representation of the system in question.

The presence of significant patterns in either or both of the residual checks indicate model inadequacy. The observed patterns that indicate model inadequacy also provide insight into the model modifications necessary to better fit the data and therefore initiate the next iteration.

Prewhitening of the input time series, although not a necessary condition of the previous two stages of the model building procedure, is a prerequisite for the diagnostic checking stage if the input series is not a white noise series.

The philosophy of this stage of the model building technique is twofold: first, to provide evidence concerning the model inadequacy and second, to provide insight into the necessary modifications required to adequately fit the data thereby initiating the next iteration, if necessary. The diagnostic checks must be sensitive to discrepancies which are likely to occur. No diagnostic checking procedure is complete in that it is possible that characteristics of an unexpected kind in the data can be overlooked. However, if a reasonably large data base is available and is analyzed using thoughtfully devised diagnostics which fail to show model inadequacies, the user should rightly feel more comfortable in applying that model. A more in-depth description of the procedure is found in Box and Jenkins [1].

2.4 Experimental Design

For identification of dynamic systems, experimental design determines the amount and type of information which will be provided by the experiment(s). The selection of input signals, sampling interval, open or closed loop identification, and on-line or off-line identification will directly influence the results of the identification technique.

Goodwin and Payne [20], Astrom and Eykhoff [6] and Isermann [5] all concur that any experimental design must take into consideration the constraints on the allowable experimental conditions. Goodwin and Payne [20] conclude by stating that the real purpose of experimental design is to maximize the information content of the data within the limits imposed by the given constraints.

Astrom and Eykhoff [6] and Isermann [5] state that the determination of the design criteria will depend heavily on the final goal of the identification problem and that examination of a priori information, final purpose of the identification and process operating conditions will determine the experimental design and the identification technique chosen due to their interdependence.

Significant simplifications in the computations can be achieved by selecting input signals of a special kind, for example step functions. But from the point of view of applications, it seems desirable to use techniques which do not make strict limitations on the input. Astrom and Eykhoff [6], Isermann [5], and Box and Jenkins [1], and MacGregor [8] all present arguments concerning the use of perturbation signals or normal operating conditions. Whenever feasible, they all support the use of perturbation signals, the choice being dependent on the identification procedure, for better system identification. For the Box-Jenkins approach both Box and Jenkins [1] and Isermann [5] recommend a pseudo random binary sequence (PRBS) input. It is an optimal input signal in that it significantly simplifies the identification procedure and, when limited a priori knowledge is available, provides a sensible initial choice of input [1].

When deciding between open- and closed-loop identification, the effects on the overall operation of the process or entire plant is of major concern, as stated previously. When open loop identification can be performed without significant disruption to the process, it is recommended since the identification approach is greatly simplified and a wider selection of identification techniques is available. Isermann [5], Astrom and Eykhoff [6], MacGregor [8], and Eykhoff [2] discuss the advantages and disadvantages of either approach in considerable detail.

Deciding between on-line or off-line identification is determined by the final goal of identification, the type of process and available computing power. In brief, if a stationary or slowly varying process is to be identified for general process analysis or for the design of control algorithms, off-line identification is usually preferred; for time varying processes and especially for adaptive control, on-line identification is considered mandatory.

For the identification of discrete-time models, the sampling time has to be chosen correctly before the experiment starts. Anderson [21] states that if a series is observed too frequently, redundant data are in fact being collected whereas, if it is recorded too rarely, high frequency detail will be lost. Isermann [5] states that a proper choice of sampling interval in most cases is not critical, because the possible range between values that are too small and those too large is relatively wide. He then presents the following widely accepted general rule of thumb,

$$T_{95}/T_{s} \approx 5 \ldots 15 \qquad (43)$$

where T_{95} is the 95% settling time of a transient function and T_{s} is the sampling rate.

Since the system in question did not have adequate stable control, process constraints imposed by the process engineers for the initial stage of the investigation were not rigid and open-loop, perturbation-input data collection was permitted.

2.5 Review of Industrial Applications of the Box-Jenkins Technique

Reported investigations of the Box-Jenkins procedure for real time system identification and control in the industrial or pilot-plant environment are very limited in the literature.

The earliest reference, by Bacon and Howe [23], appeared in 1972. The primary goal of the study was the assessment of the utility of the Box-Jenkins approach for the evaluation and modification of an industrial process. The industrial system consisted of a falling film evaporator designed to effect a separation of a monomeric material from unreacted raw materials. The authors concluded that the methodology proposed by Box and Jenkins was adaptable to the dynamics of a typical full scale chemical unit process. Furthermore, the procedure was found to be extremely useful for the development of conventional and more importantly, novel control functions that supplement existing control schemes. Bacon and Howe recommended the procedure for other chemical processes.

Box and MacGregor [24] applied the Box-Jenkins approach for analysis of closed loop dynamic-stochastic systems. The closed loop identification is achieved by using a dither signal superimposed on the system. Box and MacGregor stated that the use of a dither signal allows accurate closed-loop system identification and presented examples from the process industries. They also identified possible problems with the Box-Jenkins procedure if the stated condition of open-loop operation is assumed but is not valid.

Phadke and Wu [25] demonstrated successful implementation of a similar methodology for a multi-input-multi-output system under closed loop operation.

The usefulness was further investigated by Wright and Bacon [26] using a pilot scale heat exchanger system. They stated that statistical parametric modeling experience is limited in the literature. The probable cause, in their view, is the lack of familiarity with the underlying time series analysis techniques. Wright and Bacon concluded that the approach was a straightforward method of system identification for a noise infected dynamic single input-single output system.

Goford, MacGregor and Wright [27] presented the successful application of a predictive stochastic feedforward-feedback control of a heat exchanger stirred tank system. The underlying system identification used was the Box-Jenkins procedure.

Though limited in scope of applications, the literature presented a favorable review of the Box-Jenkins technique for system identification of industrial chemical processes. However, a need for further investigations is recognized. The scope of this thesis is to provide additional reported experience and information concerning the applicability of the Box-Jenkins approach for industrial system identification problems.

3.0 DESCRIPTION OF THE AIRFIN EXCHANGER PROBLEM

The Box-Jenkins model building technique was applied to a heat exchanger system at the Imperial Oil refinery in Sarnia, Ontario.

3.1 Process System Description

The system studied consisted of an airfin heat exchanger whose outlet stream is fed into a drum. A second process stream, originating from knock-out drums is also fed into the same drum. The system configuration is given in Figure 3. The temperature of the drum outlet stream is the system variable to be controlled. The flow rate and temperature of the inlet stream to the heat exchanger are unmeasured and uncontrolled. The flow rate and temperature of the second input stream to the drum are also unknown. However, its total flow is never more than 10% of the outlet stream from the exchanger. The known or measured system variables are louvre position, drum temperature, exchanger outlet temperature and pleneum temperature of the exchanger. The single available manipulated variable is louvre position of the heat exchanger, which is controlled pneumatically.

3.2 Problem Definition

The control objective was to maintain the drum temperature at the required setpoint. A schematic of the control strategy is illustrated in Figure 3. Deviations in

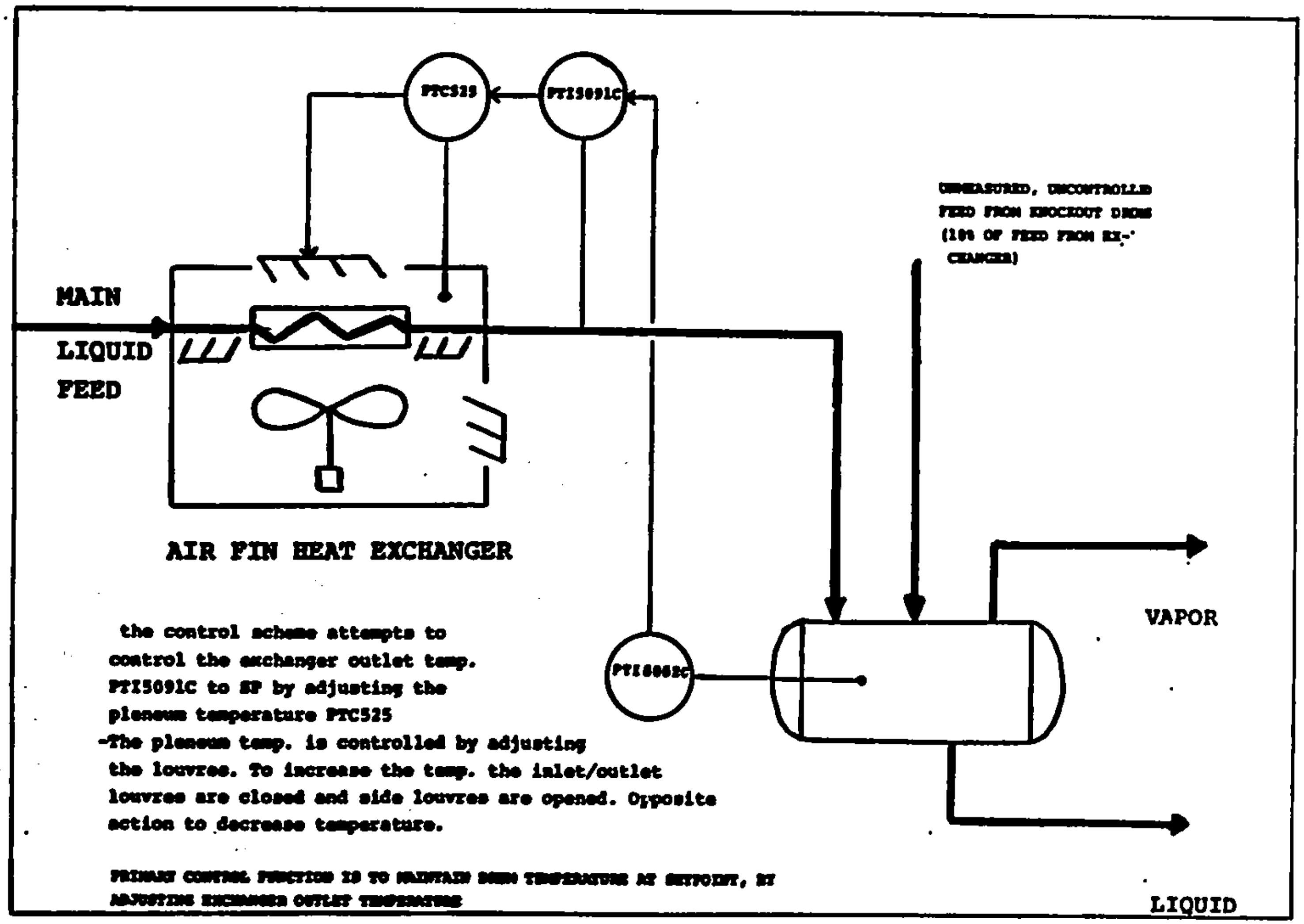


Figure 3: The Control Strategy of the Airfin Heat Exchanger

drum temperature (PTI 5052C) were corrected through manipulation of the temperature setpoint of the heat exchanger outlet stream (PTI 5091C). The exchanger outlet stream temperature (PTI 5091C) was controlled by modification of the pleneum temperature (PTC525) setpoint. Finally the pleneum temperature (PTC525) of the heat exchanger was maintained at its specified setpoint through manipulation of the pneumatically activated louvres.

This control strategy was considered inadequate by Imperial Oil. The system became unstable under closed loop operation and failed to maintain the desired drum temperature. A sample of normal closed loop operating data is plotted in Figure 4. Figure 4 demonstrates the inadequacy of the control since the measured temperatures oscillate periodically over the time of observation (over an hour). The data shown were considered to be an accurate representation of normal operating conditions.

3.3 Objectives

The primary objective was the investigation of the airfin heat exchanger control problem using the Box-Jenkins technique and the development of a satisfactory control scheme. The planned stages of analysis to achieve this objective were the following:

i) analysis of the airfin heat exchanger system,

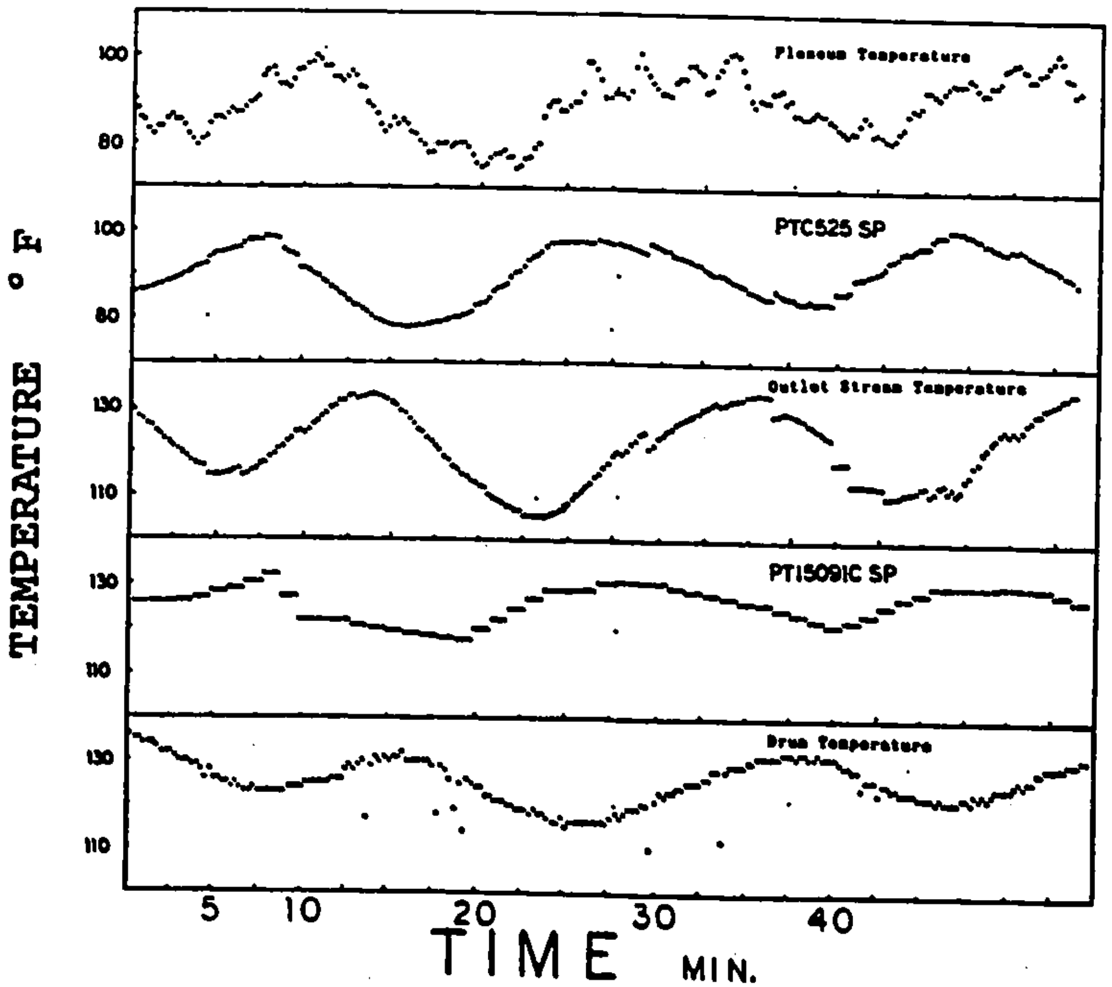


Figure 4: Normal Closed-Loop Operating Data

ii) development of an adequate model through application of the Box-Jenkins model building approach to measured system data provided by Imperial Oil; the choice of process variables restricted to those provided by the Refinery,

(iii) simulation of the airfin heat exchanger system using the transfer function-noise models developed in step (ii),

(iv) evaluation of system performance under different control strategies using the computer simulation of the airfin heat exchanger system, and

v) investigation of the possible extension to adaptive control for improved control of the system.

Due to difficulties to be described later, it was not possible to perform the simulation and control stages.

3.4 Available Data

The design of experiments for any process identification determines, in large part, the success or failure of the analysis. As discussed previously, important parameters, such as sampling interval, must be selected with great care to ensure an accurate representation of the system under study.

Unfortunately, in this case, the experimental design was developed and determined exclusively by Imperial Oil.

The author was afforded no opportunity for input into the experimental design process. The data was collected and forwarded to the university by the Refinery. The following data were provided by Imperial Oil:

i) normal open loop data,

ii) normal closed loop data,

(iii) open loop operation with a PRBS input signal applied to the pneumatic actuator of the louvres,

iv) pseudo open loop operation with a PRBS input applied to the setpoint of the pleneum temperature controller with all other control loops inactive.

The sampling rate and control interval were preset at 15 seconds by the Refinery. The impact of these aspects of the experimental design is discussed in a subsequent section.

The initial investigation of the process system produced seemingly unrealistic results which were strongly contested by Imperial Oil. The author was subsequently provided with another data set obtained using a step-function perturbation to the pneumatic actuator of the louvres during open-loop operation.

The identification and parameter estimation of the system were repeated with this data set. Results from both analyses are presented. Discussion of the inconsistencies of the results follows.

4.0 ANALYSIS OF THE PRBS OPEN LOOP DATA

The Box Jenkins model building approach was used to determine the combined transfer-function noise models relating the known system temperatures to the lone available manipulated variable, percentage of operating pressure to the louvres.

4.1 Selection of Data for Initial Investigation

A PRBS input perturbation to the system, as outlined in Section 2.4, during open loop operation, greatly reduces the computational effort and is considered the optimal choice of input signal for the Box Jenkins approach when available.

Since data collected under these conditions were made available by Imperial Oil, the initial system analysis and identification were performed using this data set. Plots of the measured system temperatures and manipulated variables are provided in Figures 5, 6, 7, and 8.

A preliminary qualitative review of the data as plotted in Figures 5 - 8 raised some concern about the validity of the data or more specifically the stated conditions of a random input and open-loop data collection. The cyclical or sinusoidal behaviour exhibited by the three system output temperatures is unexpected if the system input is actually a random process and data collection was performed under open-loop conditions. An autocorrelation

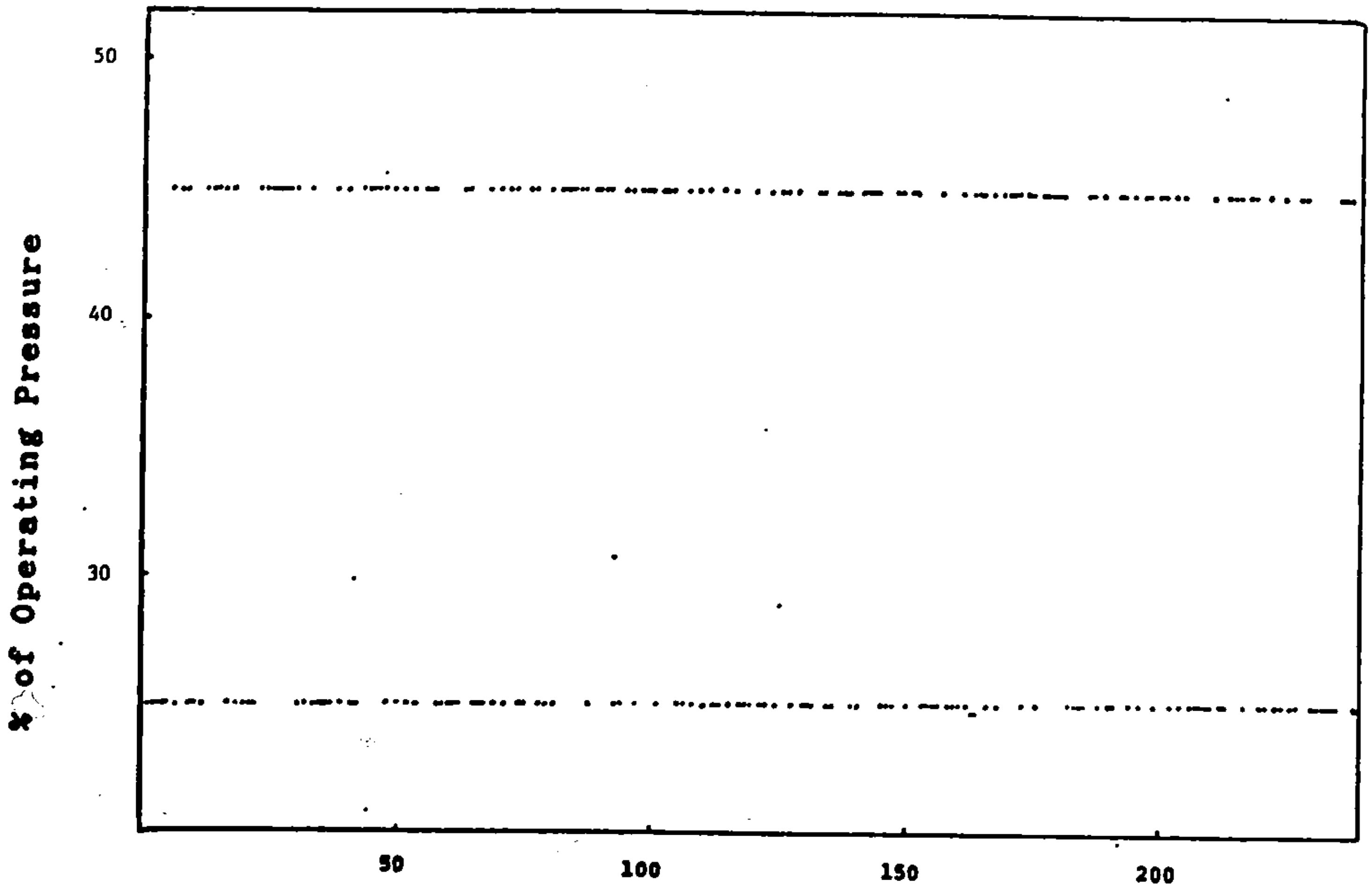


Figure 5: PRBS Data: % of Operating Pressure (System Input)

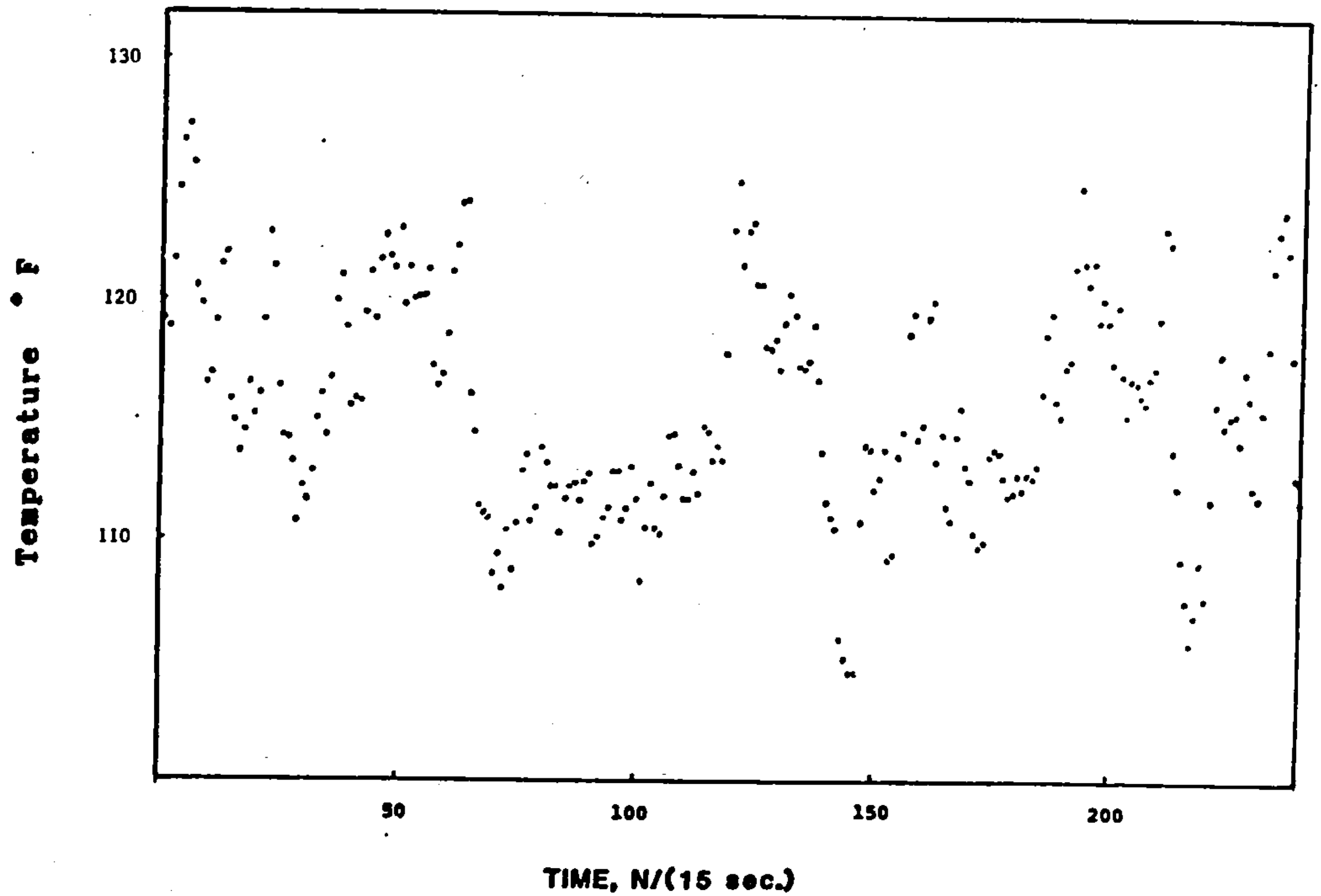


Figure 6: PRBS Data: Pleneum Temperature

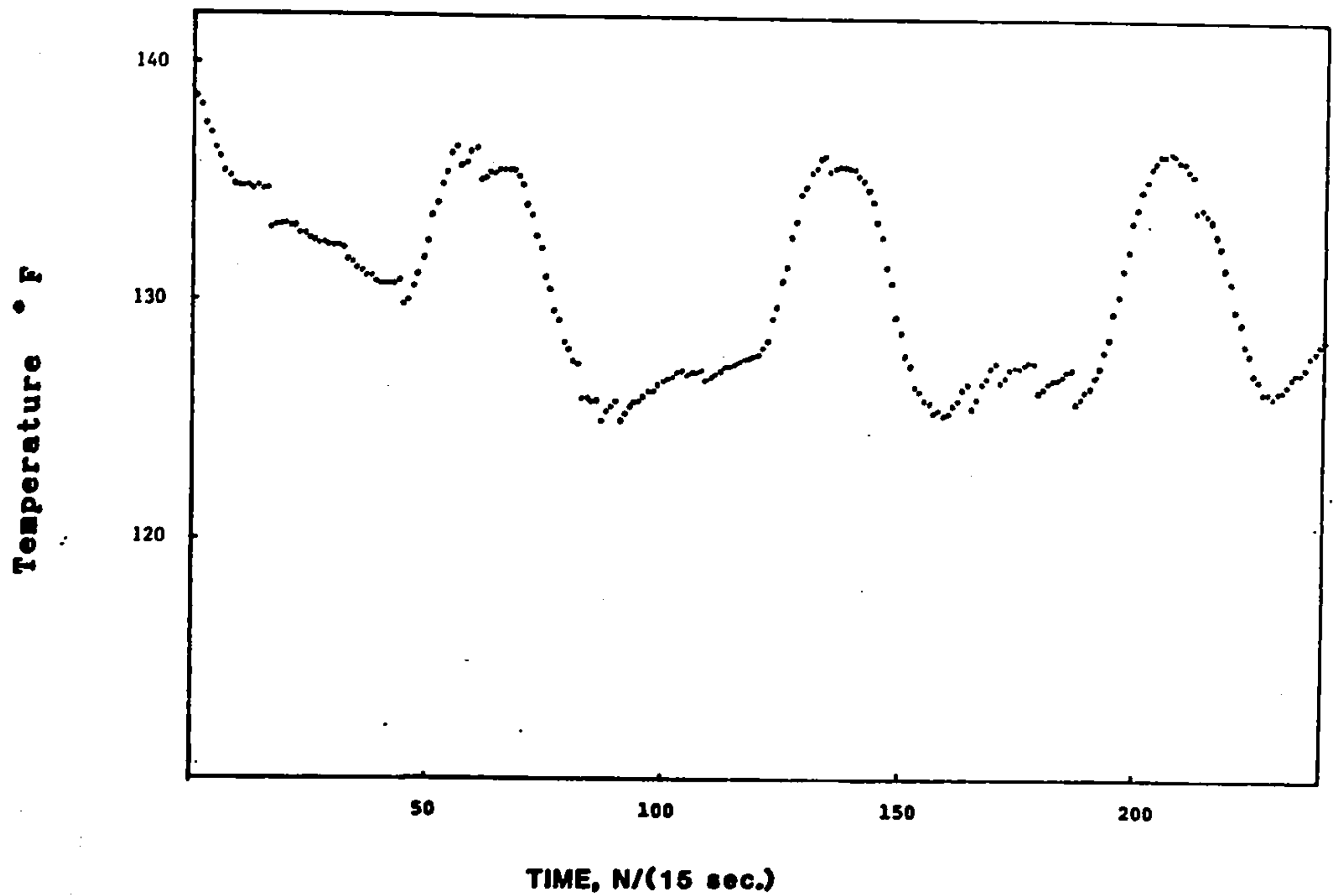


Figure 7: PRBS Data: Outlet Stream Temperature

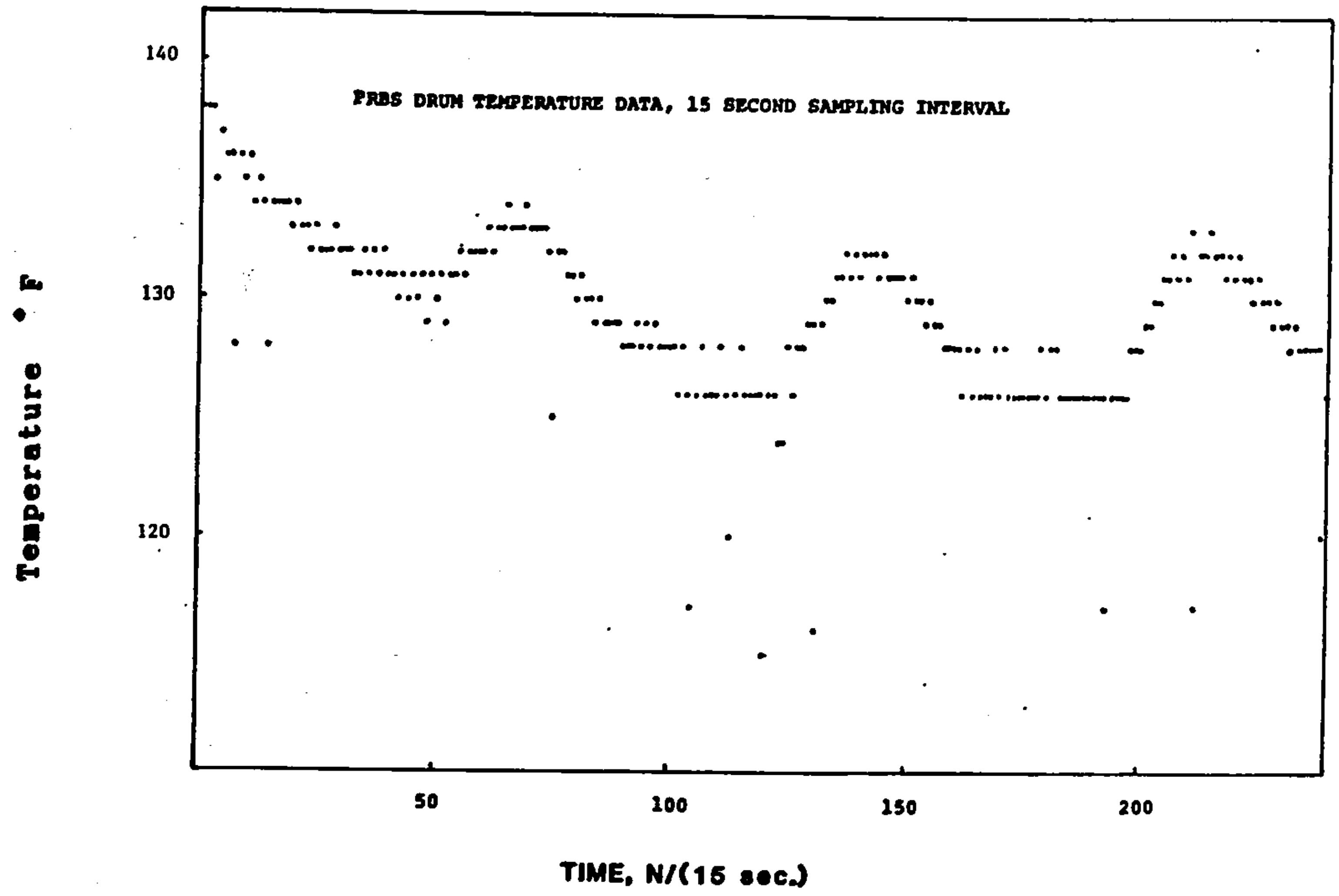


Figure 8: PRBS Data: Drum Temperature

check of the system input, discussed in more detail in the next section, verified that the input series was indeed a PRBS input as stated by Imperial Oil.

The condition of open-loop operation for the PRBS data was questioned due to the similarities of the observed temperature responses in Figures 6 - 8 when compared to the temperature responses exhibited under closed-loop operation shown in Figures 24-27.

With these reservations, the Box-Jenkins model building technique was applied using the PRBS data provided.

4.2 Preliminary Identification

The first step of the preliminary identification analysis was to verify that the PRBS input time series could be treated as white noise. The autocorrelation function of the manipulated variable was estimated and plotted in Figure 9 with its associated error bounds represented by the broken lines. By definition, a white noise series is a series of random shocks having a normal distribution with zero mean and constant variance σ_a^2. The autocorrelation of a white noise series will be zero for all positive lags, $k>0$.

In Figure 9, all values of the autocorrelation function for $k>0$ fall within the error limits. Therefore, the manipulated variable for the PRBS open loop data set may be considered as white noise.

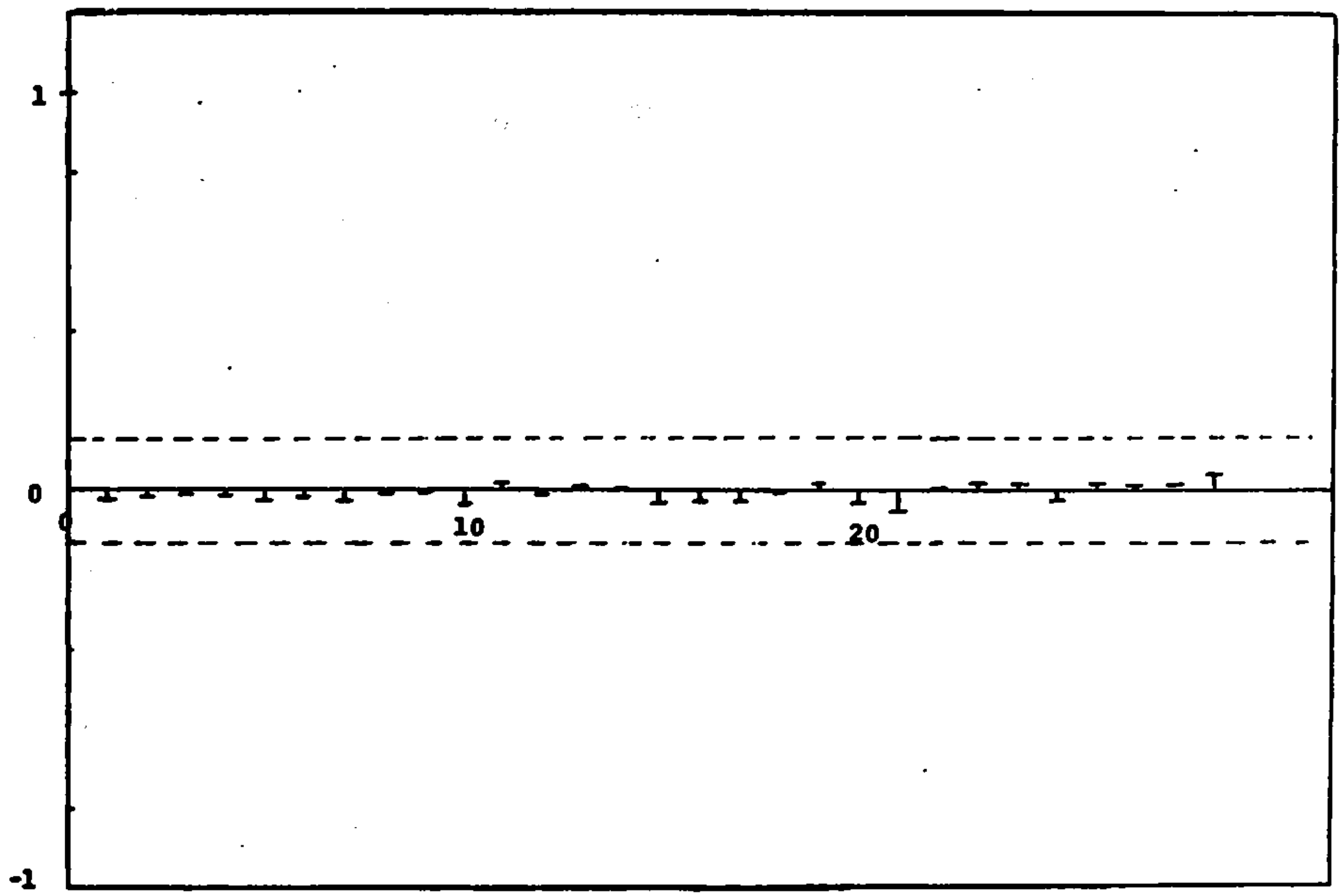


Figure 9: Autocorrelation function of the PRBS Data System Input

Estimates of the crosscorrelation functions relating the three measured system temperatures (pleneum, exchanger outlet stream and drum) to the manipulated variable were calculated and plotted respectively in Figures 10, 11, and 12. The approximate error bounds for the estimated crosscorrelation functions, at the 95% confidence level, appear as broken lines in the plots. Statistical independence between the time series is concluded should all estimates of the crosscorrelation function not be significantly different from zero.

The crosscorrelation function of Figure 10 suggested a model having the following form:

$$(1-\delta_1 B)Y_t = \omega_0 B^b X_t$$

Prewhitening renders the crosscorrelation function estimate of the input and output variables directly proportional to the impulse response function. Therefore by using the estimated values of the crosscorrelation function which are considered to be significantly different from zero and the procedure outlined in Box and Jenkins [1], the following initial estimates of the parameters were obtained.

(i) $\delta_1 = 0.60$ (ii) $\omega_0 = -0.18$ (iii) $b = 2$

All values of the crosscorrelation function for the outlet stream temperature and manipulated variable time series fell within the error bounds as shown in Figure 11.

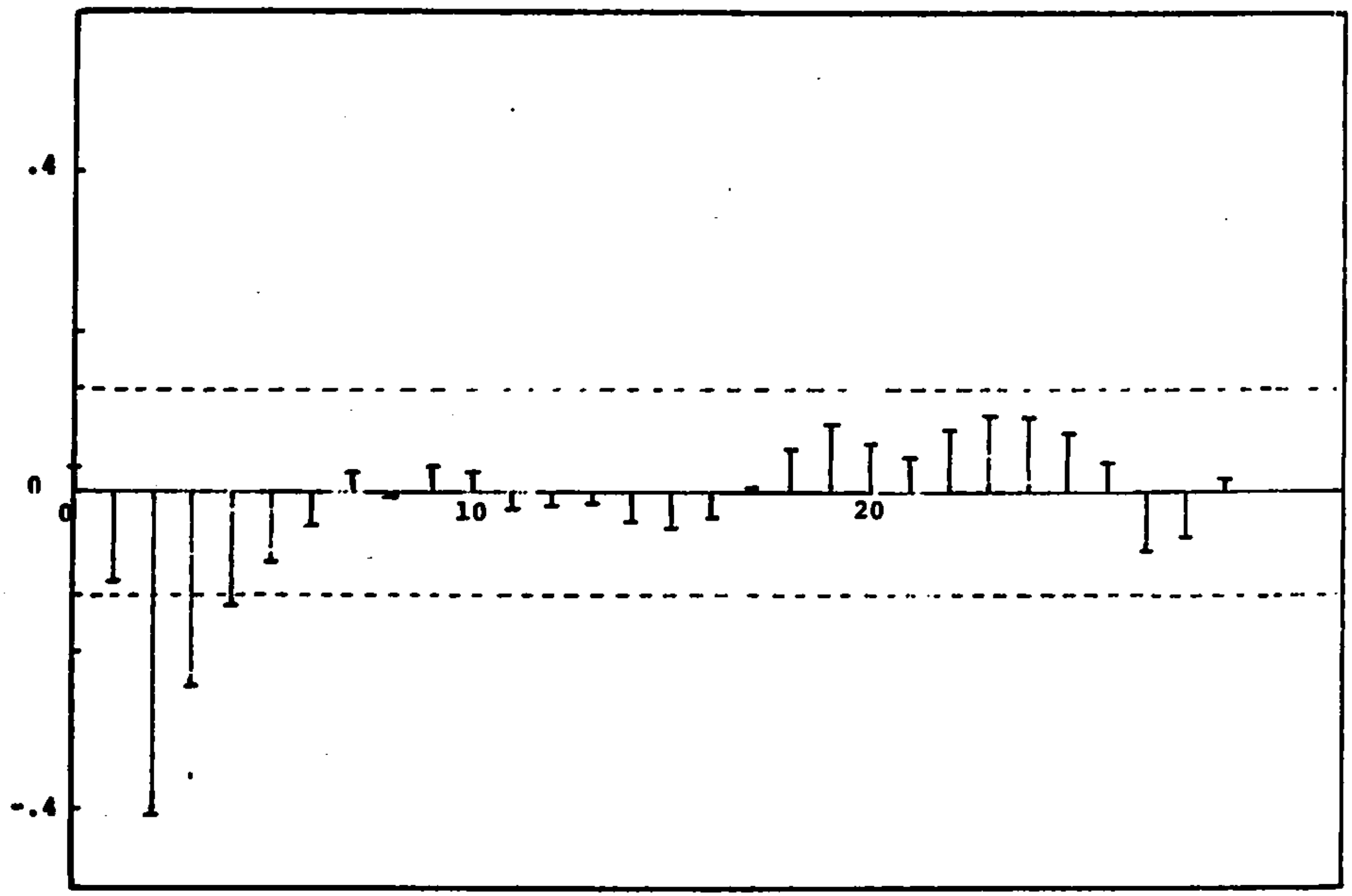


Figure 10: Crosscorrelation Function Between the PRBS Data Pleneum Temperature and System Input

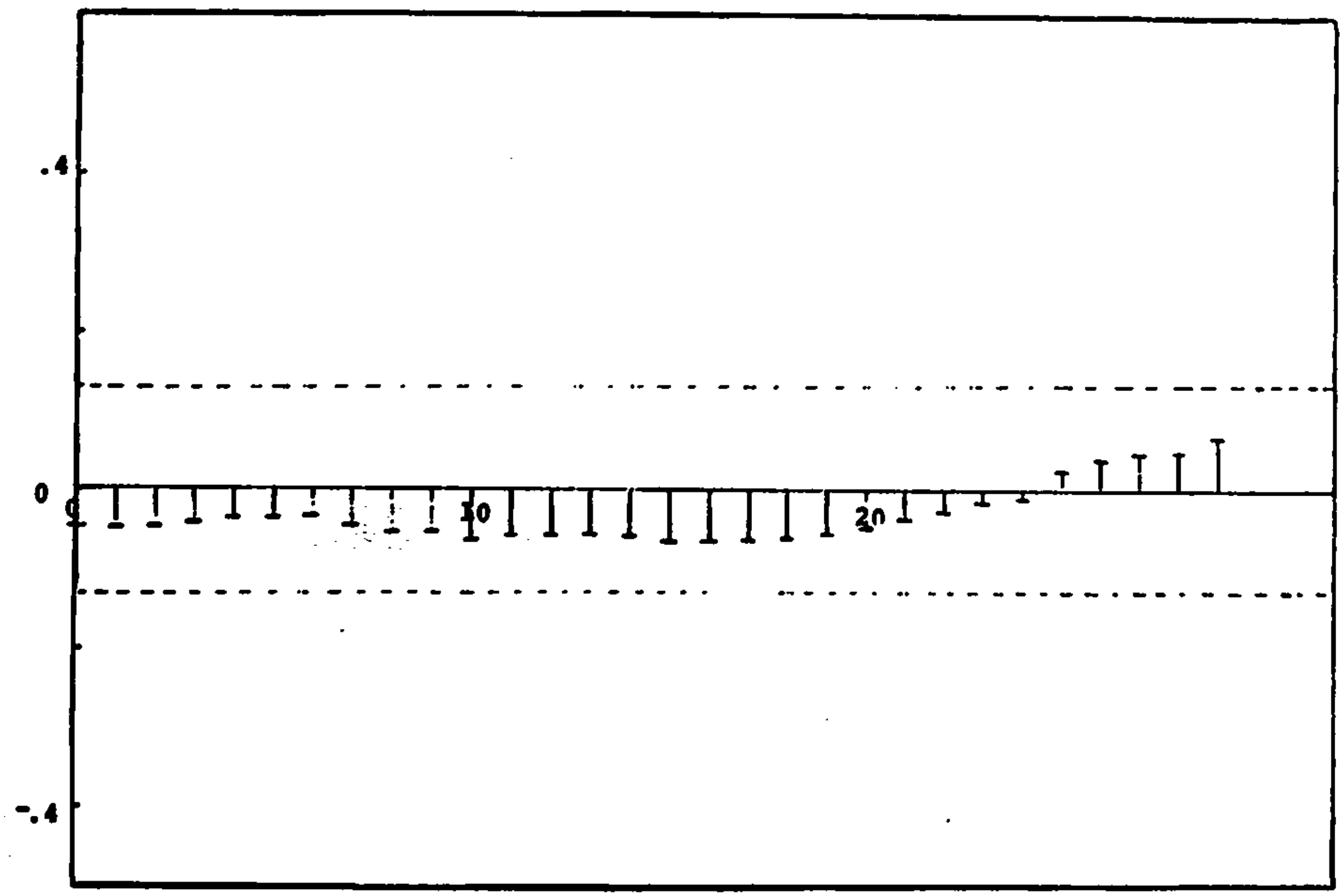


Figure 11: Crosscorrelation Function Between the PRBS Data Outlet Stream Temperature and System Input

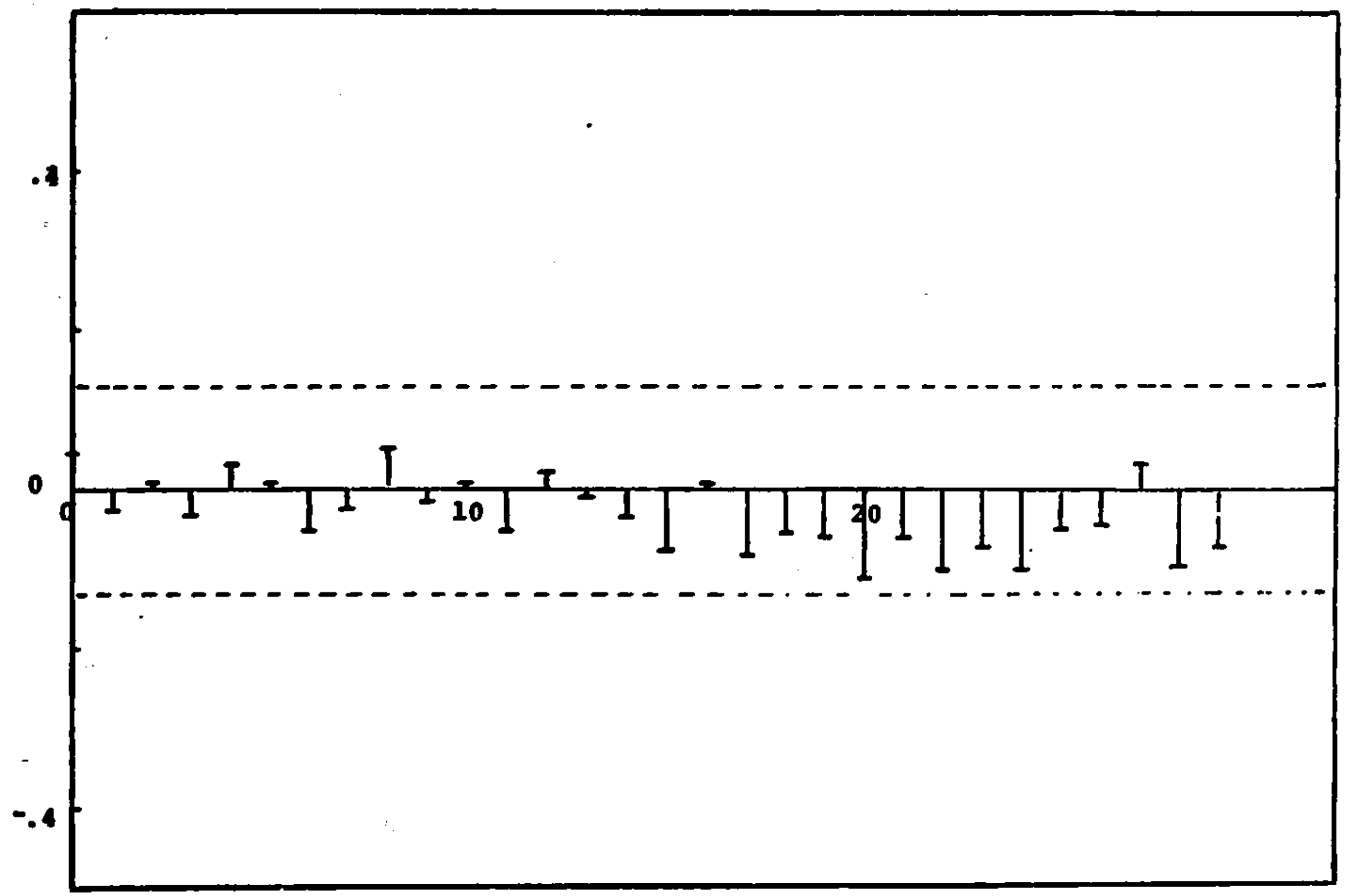


Figure 12: Crosscorrelation Function Between the PRBS Data Drum Temperature and System Input

These results indicated that no statistically significant dependence existed between the two variables.

The same results are reflected in Figure 12 for the drum temperature and manipulated variable crosscorrelation function estimates. No relationship between the two variables is again indicated.

4.3 Conclusions of the Preliminary Identification Stage

Analysis of the PRBS input data set indicated that both the heat exchanger outlet stream temperature and drum temperature were statistically independent of the lone manipulated variable. That is to say that no significant dependence exists between those two system temperature variables and the manipulated variable.

These results indicated that adequate control of the system drum temperature is not feasible using the current manipulated variable. Furthermore, the indicated relationships or lack thereof aborts any attempts of an adequate system model representation with the current measured system variables.

Interpretation of these results relative to the physical reality of the system provoked suspicion concerning the validity of the results. These conclusions indicated that outlet stream temperature and drum temperature were independent of the pleneum temperature. This arose from the observation that a change in pleneum temperature was related to a change in louvre position, but that the change was not reflected in the outlet stream of the heat exchanger. This disagreed with the basic premise of the airfin heat exchanger. The lack of confidence in the results and our suspicions were corroborated by Imperial Oil.

5.0 ANALYSIS OF THE STEP RESPONSE DATA

As a result of the unlikely conclusions drawn from the PRBS data, Imperial Oil provided additional data, consisting of a step function input applied to the louvre controller during open loop operation and sampled at one minute intervals.

The temperature responses, are shown in Figures 13, 14, and 15. Qualitative assessment indicated a significant correlation or dependence between the manipulated variable and all three temperature responses (pleneum, outlet and drum). The observed behaviours of the outlet and drum temperature responses for the step response data contradicted results presented in section 4 using the PRBS data. The discrepancies in results will be discussed in section 6.

5.1 Preliminary Identification

The step function is not a white noise series therefore determination of the pre-whitening operator for the input series is required.

The calculated autocorrelation function for the step input is shown in Figure 16. As expected this input series was found to be nonstationary, the autocorrelations did not become insignificant at large values of the lag, k. As stated previously homogeneous nonstationarities can be removed by suitable differencing (in general $d < 3$ times) of

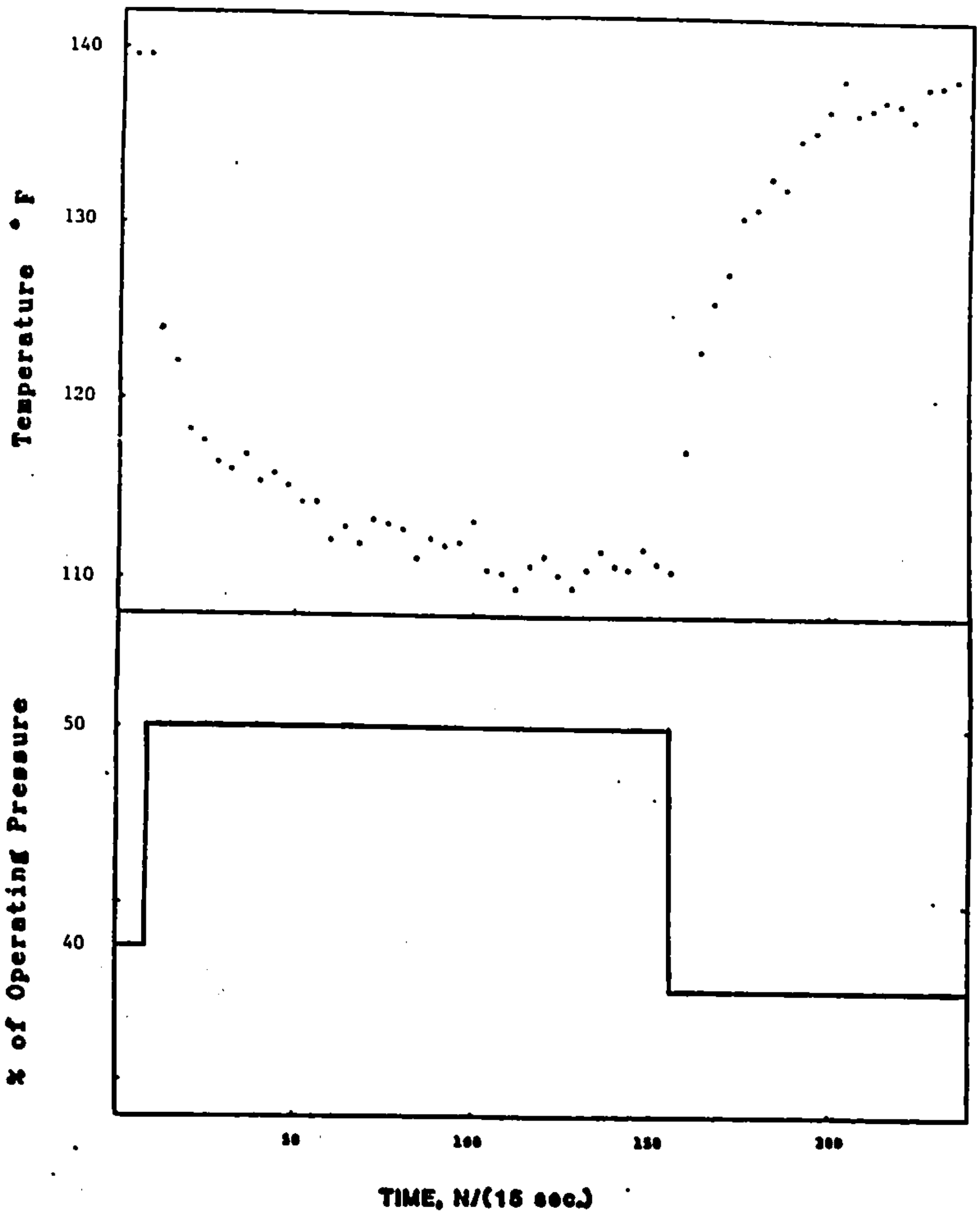


Figure 13: Step Function Data: Pleneum Temperature

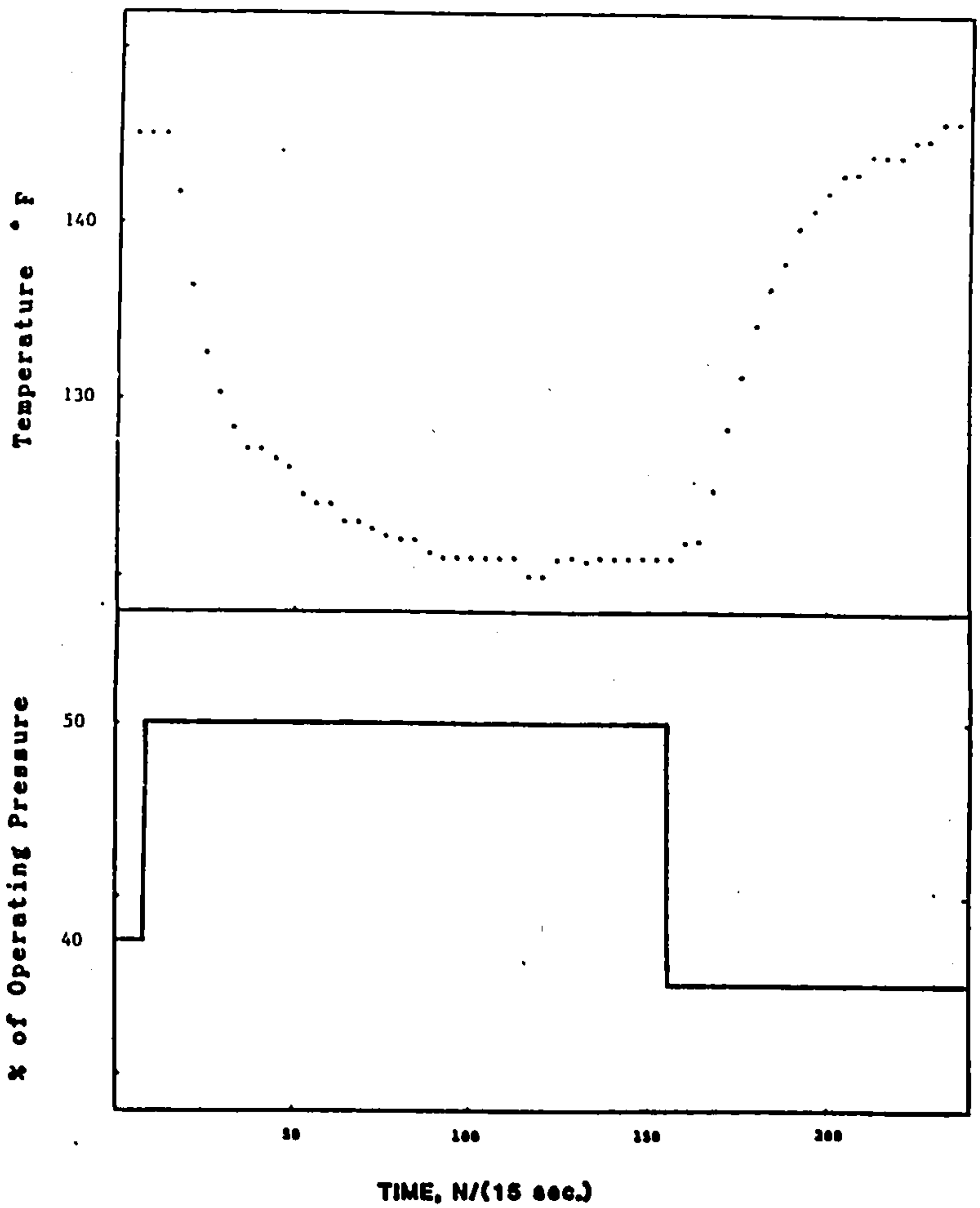


Figure 14: Step Function Data: Outlet Temperature

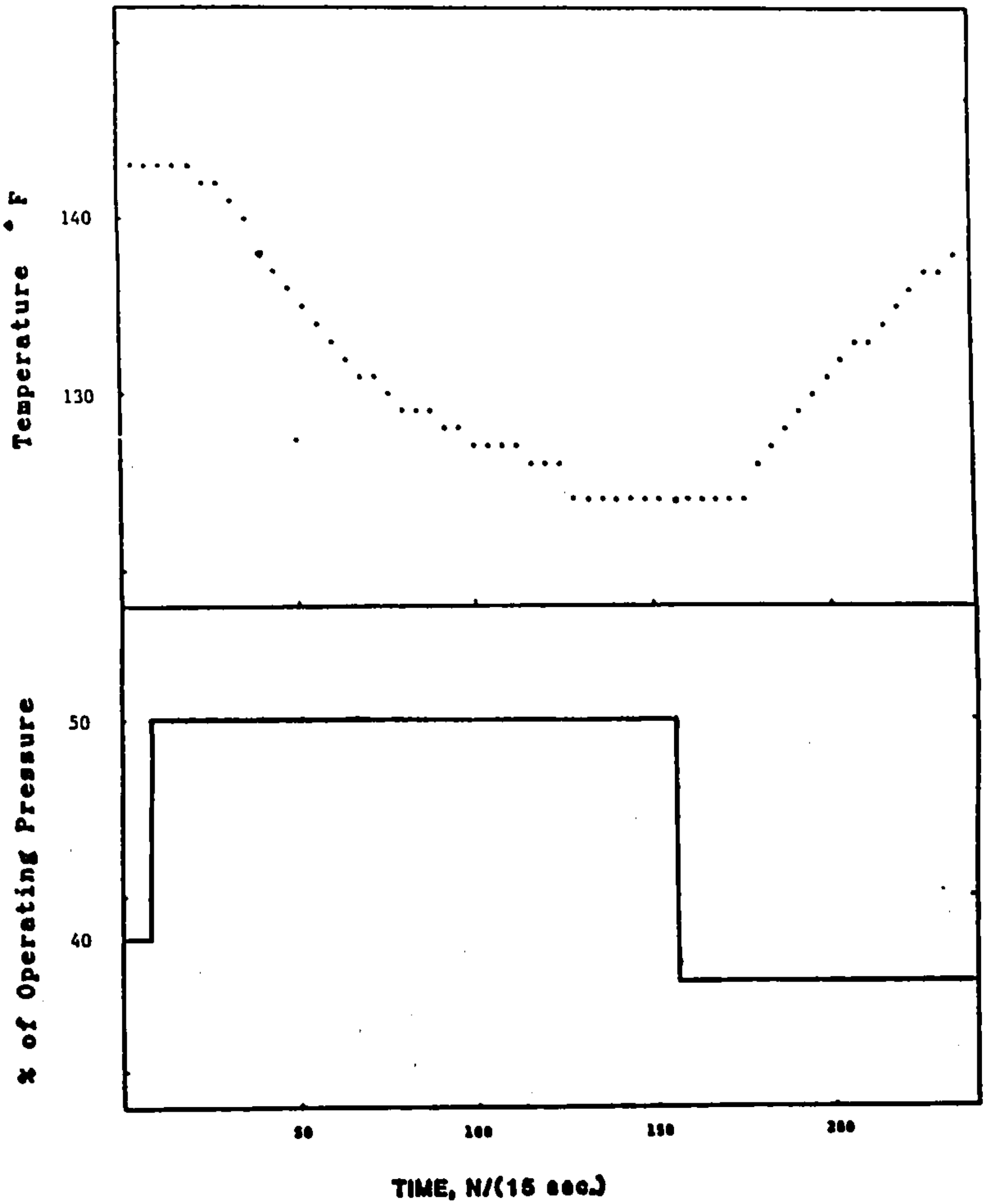


Figure 15: Step Function Data: Drum Temperature

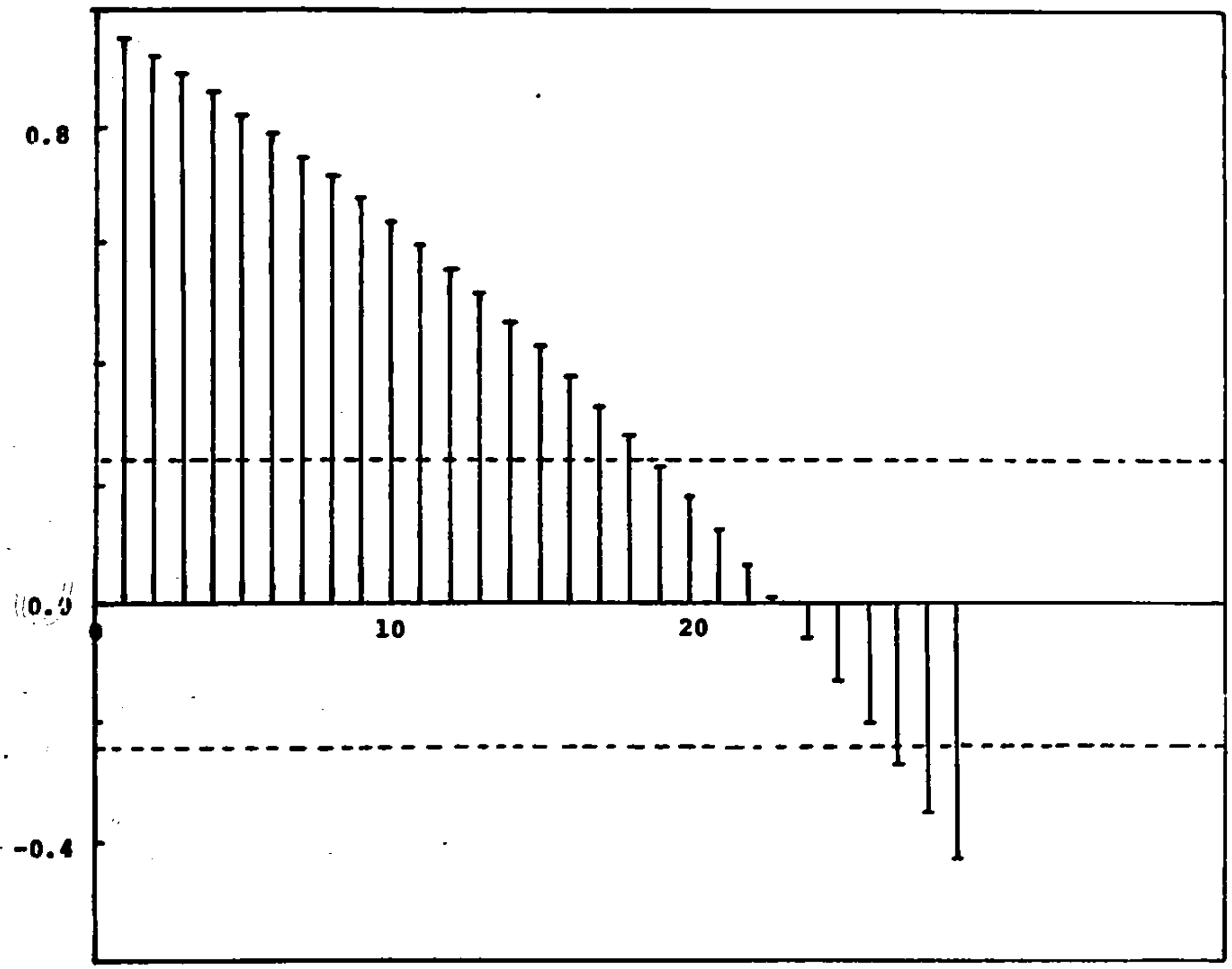


Figure 16: Autocorrelation Function of the Step Function System Input

the time series [22]. Thus, the first differences of the input were determined; the autocorrelation function was calculated and plotted in Figure 17. Stationarity was achieved and yielded a white noise series. The appropriate prewhitening operator is therefore simply the first degree difference operator $(1-B)$.

The prewhitening operator was then applied to the input time series and the three system temperatures. The crosscorrelations between the transformed temperatures and the prewhitened input were calculated and are shown in Figures 18, 19, and 20.

The preliminary identification criteria of the Box-Jenkin procedure was then applied to the estimated crosscorrelation functions of the three bivariate processes. The model form(s) initially suggested for each relationship follow in table 1.

5.2 Estimation of the Transfer Function-Noise Model

The iterative estimation and diagnostic checking stages of the Box-Jenkins technique, as outlined in section 2.3.2 were then applied to each of the proposed models for the three bivariate stochastic pairings. Using the Box-Jenkins criteria for model adequacy, iterations were performed until adequate transfer function-noise models for each of the three bivariate stochastic processes were determined. The final transfer function-noise model

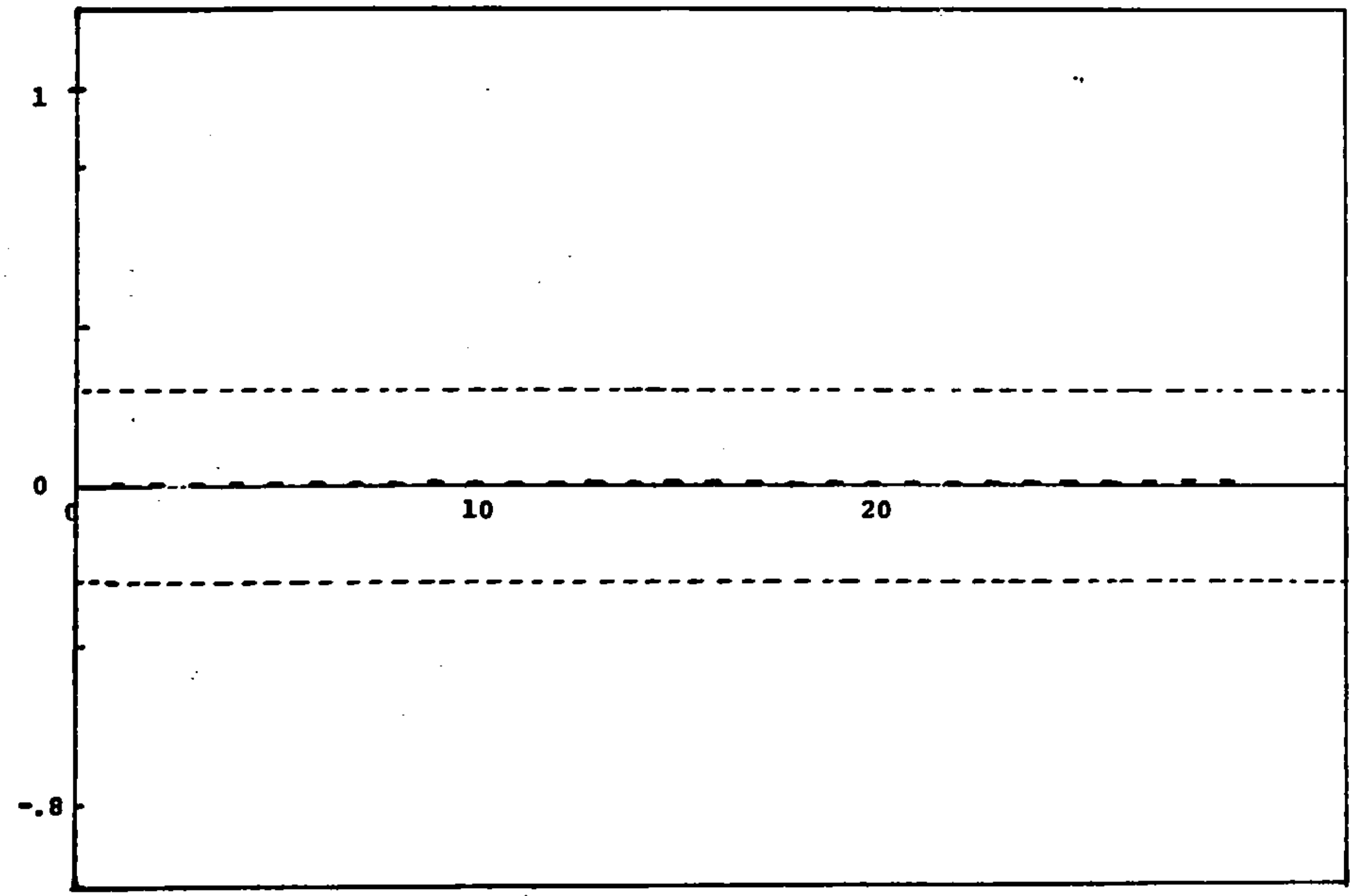


Figure 17: Autocorrelation Function of the Differenced (1st degree) Step Function System Input

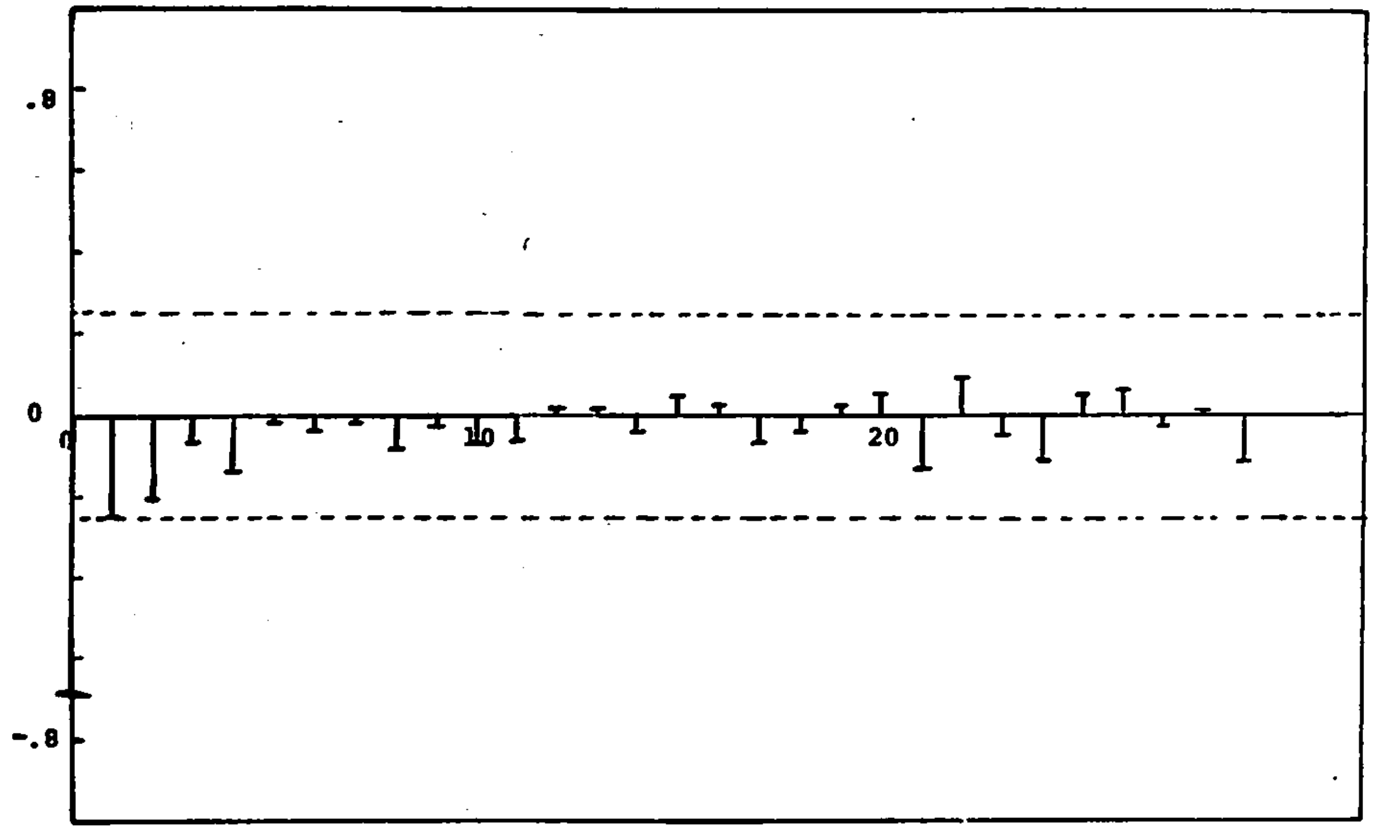


Figure 18: Crosscorrelation Function Between the Step Function Transformed Pleneum Temperature and Prewhitened System Input

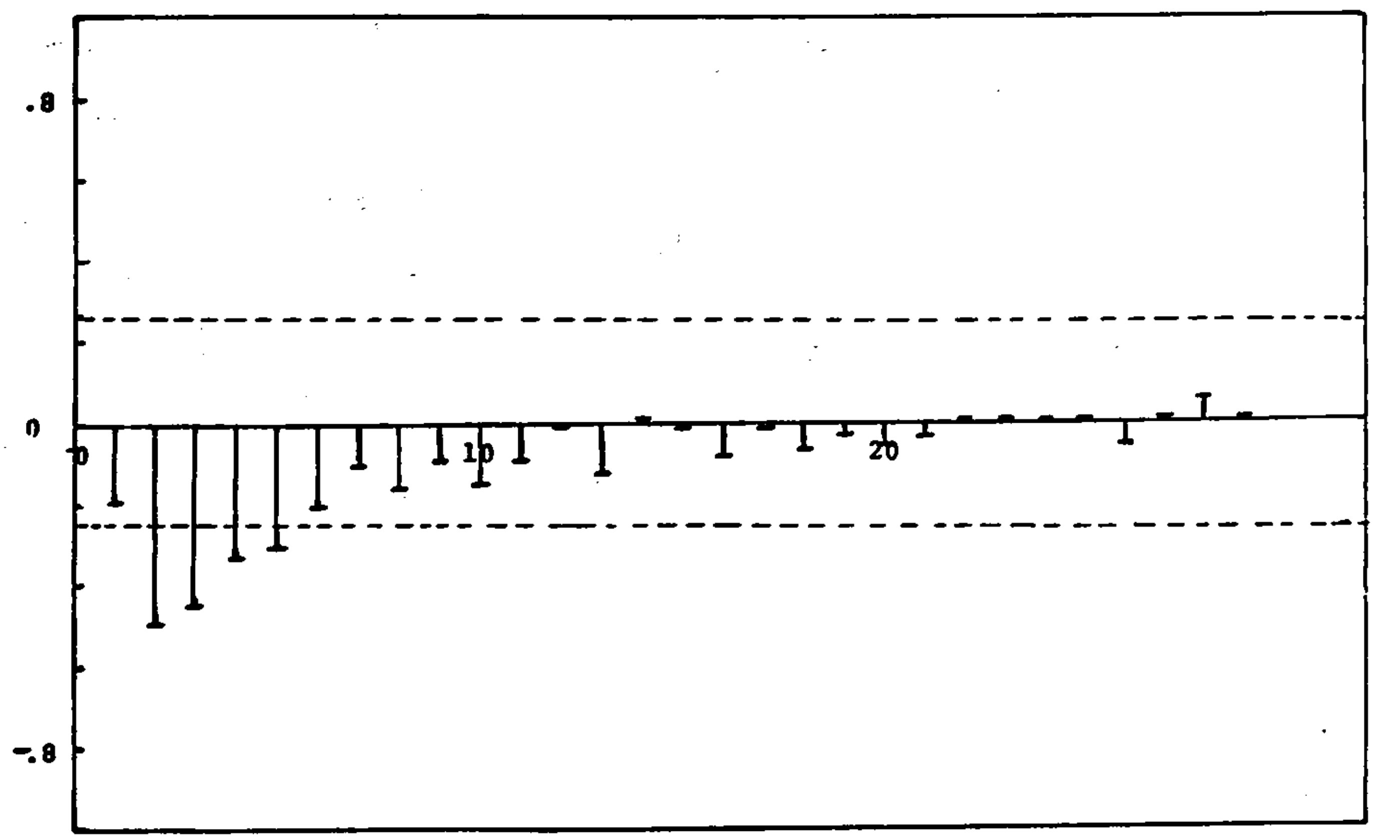


Figure 19: Crosscorrelation Function Between the Step Function Transformed Outlet Temperature and Prewhitened System Input

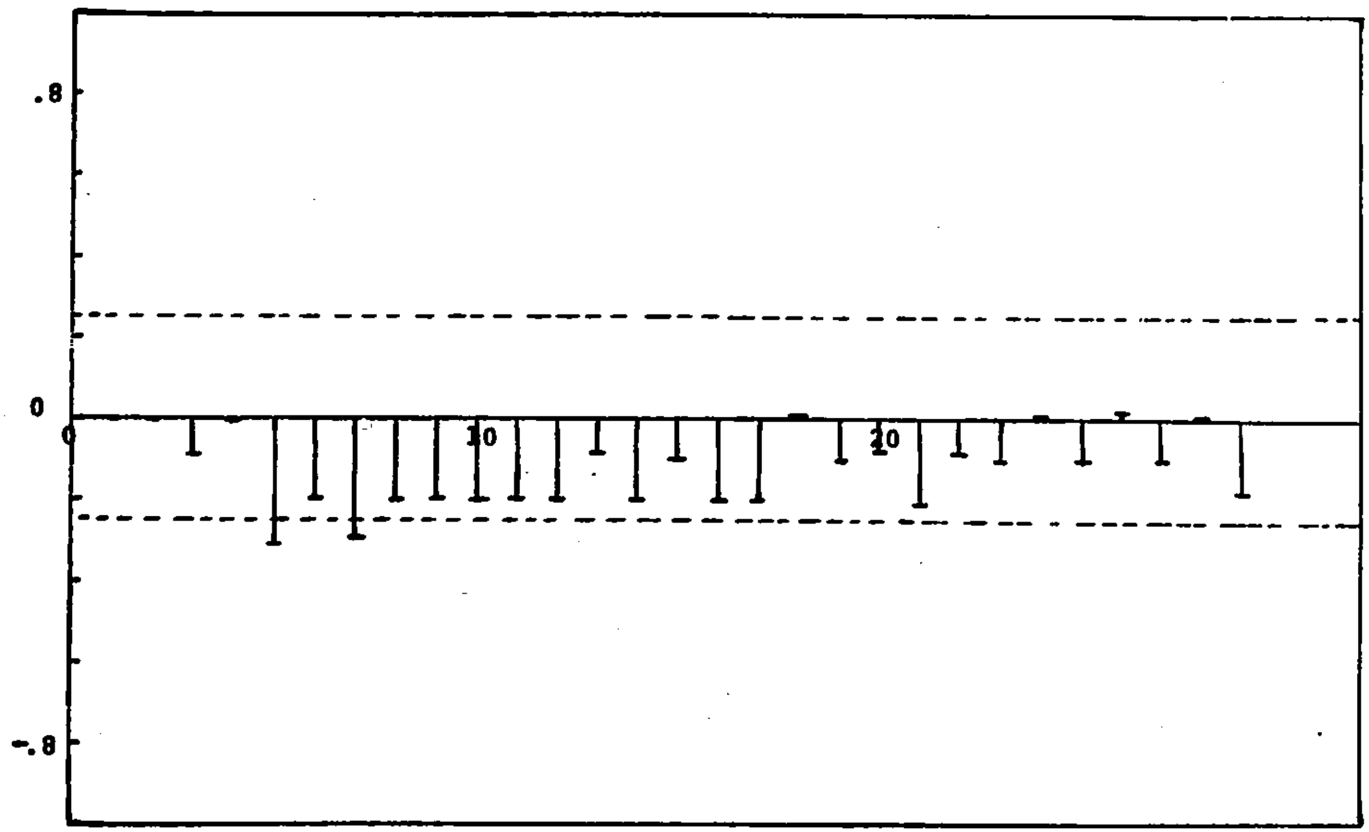


Figure 20: Crosscorrelation Function Between the Step Function Transformed Drum Temperature and Prewhitened System Input

Table 1: Initial Transfer Function Model For The Airfin Heat Exchanger System

Process Temperature	Identified Model(s)
Pleneum Temp.	$\nabla Y_t = \omega_0 \ \nabla X_t$ $\nabla Y_t = (\omega_0 - \omega_1 B) \ \nabla X_t$
Outlet Temp.	$(1-\delta_1 B) \ \nabla Y_t = \omega_0 \ \nabla X_{t-2}$
Drum Temp.	$(1-\delta_1 B) \ \nabla Y_t = (\omega_0 - \omega_1 B) \ \nabla X_{t-5}$ $(1-\delta_1 B-\delta_2 B^2) \ \nabla Y_t = \omega_0 \ \nabla X_{t-5}$

representations for each of the three bivariate processes are shown in Table 2.

The results of the quantitative analysis of the step response data demonstrate a significant correlation between all three system temperature variables and the manipulated system input. Furthermore, the choice of manipulated variable is shown to be reasonable since a 15% deviation in the manipulated variable produces a significant output response for all three temperatures.

5.3 Conclusions

The analysis of the step response data provided by Imperial Oil determined that significant correlation exists between the manipulated system input and all three system temperature variables. These results are inconsistant with the conclusions drawn from the analysis of the PRBS data. Discussion of the contradictory results follows.

Table 2: Combined Transfer Function–Noise Models for the Airfin Heat Exchanger System

Input/ Output	Transfer Function	Noise Model
∇% op. press ∇ Pleneum T	$(1-0.70B)Y_t = -0.30X_{t-1}$	$\nabla N_t = a_t$
∇% op. press ∇ Outlet T	$(1-0.80B)Y_t = -0.30X_{t-2}$	$\nabla N_t = a_t$
∇% op. press ∇ Drum T	$(1-0.88B-0.05B^2)Y_t = -0.11X_{t-5}$	$\nabla N_t = MA(1)$ process

6.0 DISCUSSION OF RESULTS

Striking inconsistencies existed between the results obtained from the analyses of the different data sets presented in the previous sections. For meaningful conclusions to be drawn from this research, an explanation of the observed discrepancies was necessary. Examination of both the estimation procedure and the process operation lead to four possible factors, which, taken separately or together, could have possibly accounted for the observed disagreements in results. These factors were:

(i) The data was inappropriate.

(ii) Process parameters changed between data collection periods.

(iii) The computer algorithm used to perform the Box-Jenkins identification procedure was inaccurate.

(iv) The Box-Jenkins identification procedure is not an adequate technique for use in industrial system identification.

Each of the possible causes for inconsistencies in the results was investigated and the findings are presented in the subsequent sections.

6.1 Inadequate Data

A possible cause for the contradictory results may be inappropriate data. The purpose of the measured data is to provide an accurate picture of the system under study.

The step function open loop data are plotted in Figures 13, 14, and 15. Significant dependence between the manipulated variable and the measured system temperatures is assumed to be true. The observed responses indicated a first or second order plus dead time model for the three bivariate processes.

The system, under the stated condition of normal open-loop operating conditions, is represented in Figures 21, 22, and 23. The airfin heat exchanger system was switched from normal closed-loop operation to open-loop operation by opening all control loops and setting the manipulated variable to a fixed value. The transition from closed-loop operation to open-loop operation is comparable to application of a step function to the manipulated variable although the system is not at steady state prior to the deterministic perturbation. This allows for the comparison of the two data sets.

Steady state is achieved for all three system temperatures within a maximum of approximately 25 minutes for both data sets. The individual temperature responses are consistent for both data sets although their relative deviation is significantly greater for the step function open-loop data set.

The PRBS open-loop data, plotted in Figures 7, 8, 9, and 10, display a similar response pattern for the three measured system temperatures. A sinusoidal function with an

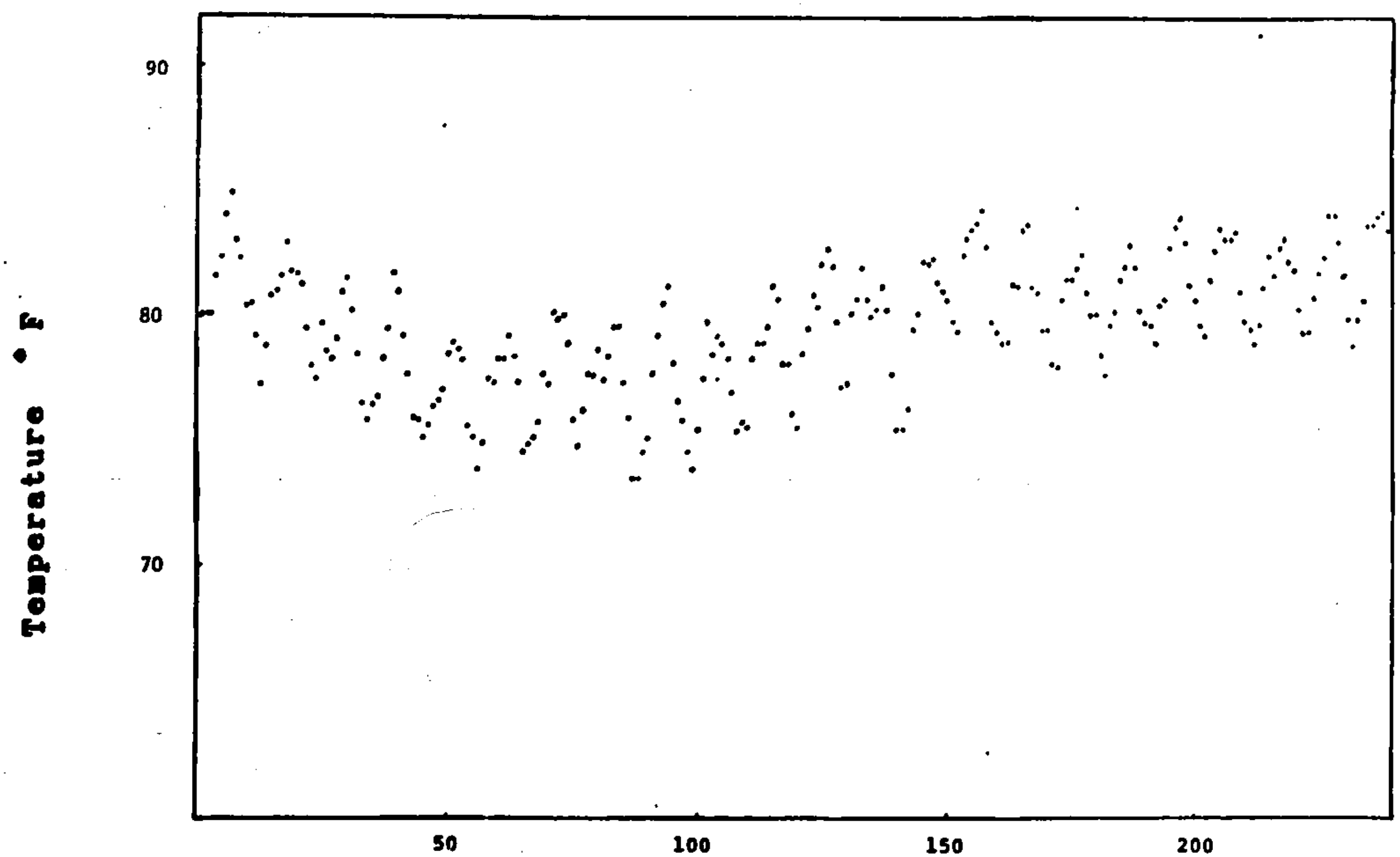


Figure 21: Normal Open-Loop Operating Data: Pleneum Temperature

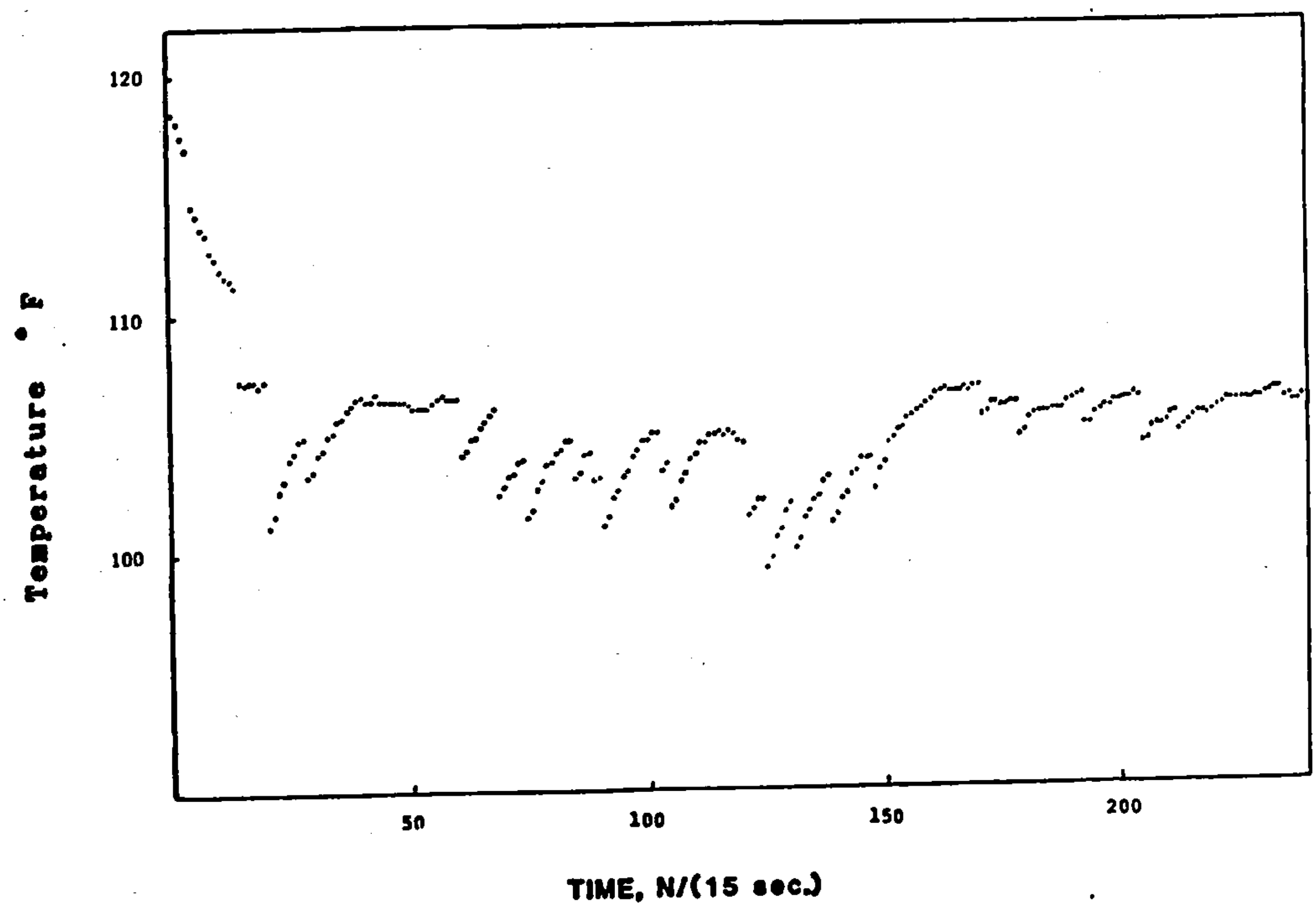


Figure 22: Normal Open-Loop Operating Data: Outlet Temperature

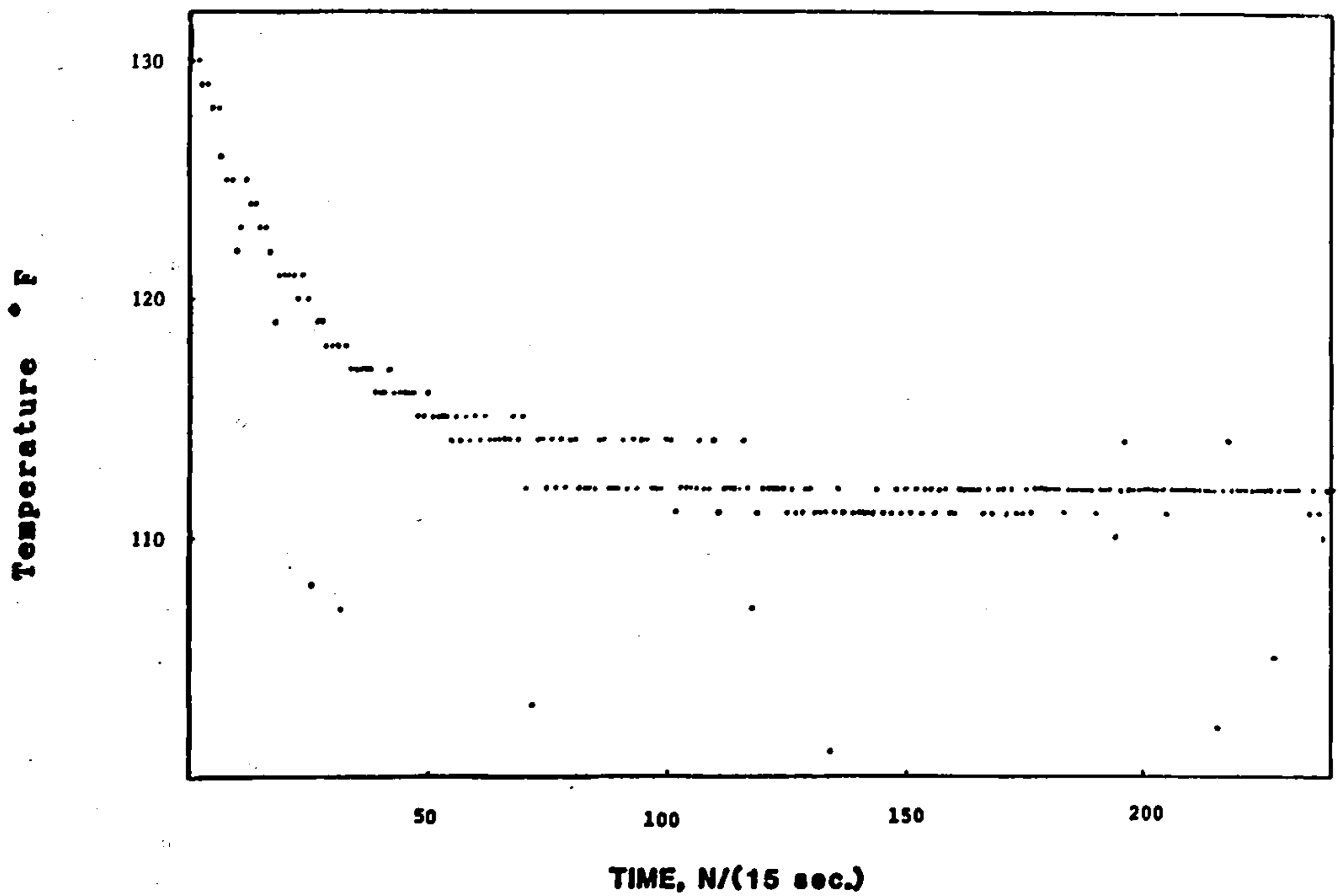


Figure 23: Normal Open-Loop Operating Data: Drum Temperature

approximate period of 20 minutes is observed. The relative amplitude of the sine wave diminishes between temperature variables, the pleneum temperature possessing the greatest amplitude and the drum temperature the smallest. This is consistent with the observed system behaviour demonstrated by the other open-loop data sets. Reaction to system disturbances is dampened along the system. Furthermore, the observed delay time also increases along the system, the pleneum temperature having the shortest delay, the drum temperature the largest.

In summary, the observed time delays and amplitude response dampening for the PRBS data were consistent and expected based on the normal open-loop and step function open-loop data sets. However, the periodic wave form superimposed upon the system responses was unexpected.

Finally, the normal closed-loop operation data are presented in Figures 24, 25, 26, and 27. An overwhelming similarity of the individual system temperature responses as stated previously is observed for the PRBS open-loop and normal closed-loop data sets. The closed-loop data response for all three system temperatures exhibit the same sinusoidal wave form with approximate period of 20 minutes. This behaviour is identical to that observed for the PRBS open-loop data. Similar characteristics of temperature response amplitude and time delays are also detected for the two data sets. However, the manipulated variable behaviour

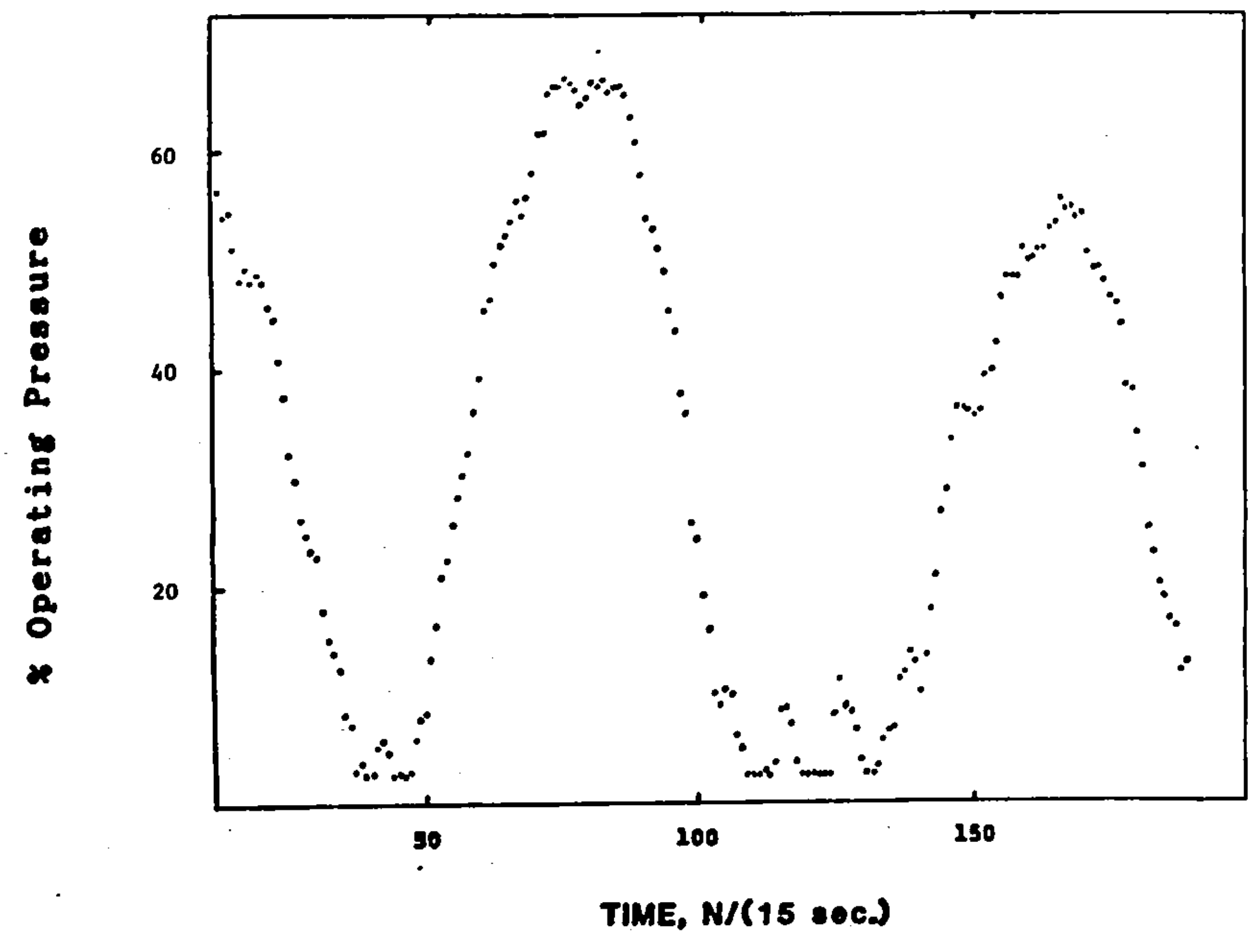


Figure 24: Normal Closed-Loop Operating Data: % Operating Pressure

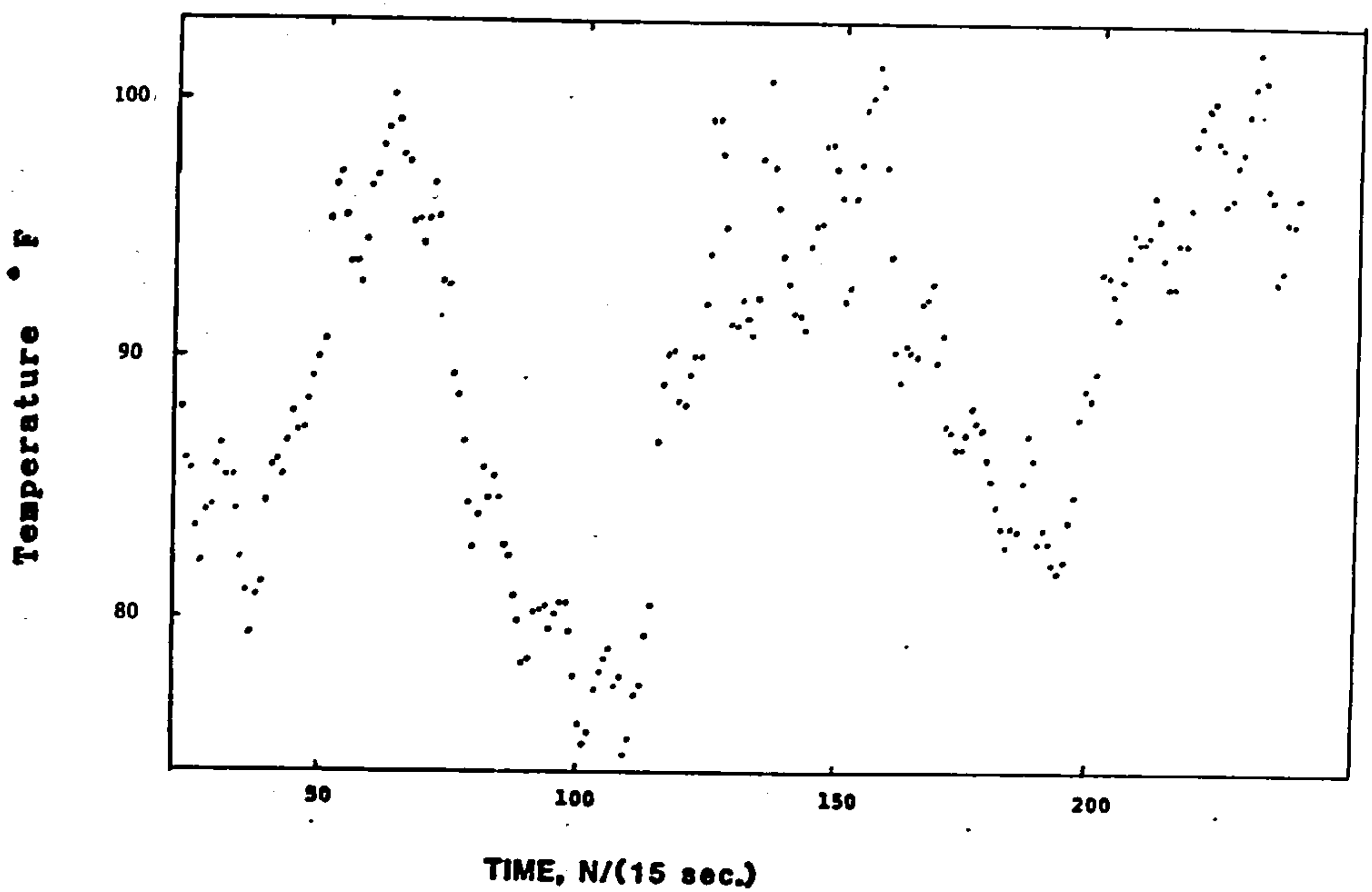


Figure 25: Normal Closed-Loop Operating Data: Pleneum Temperature

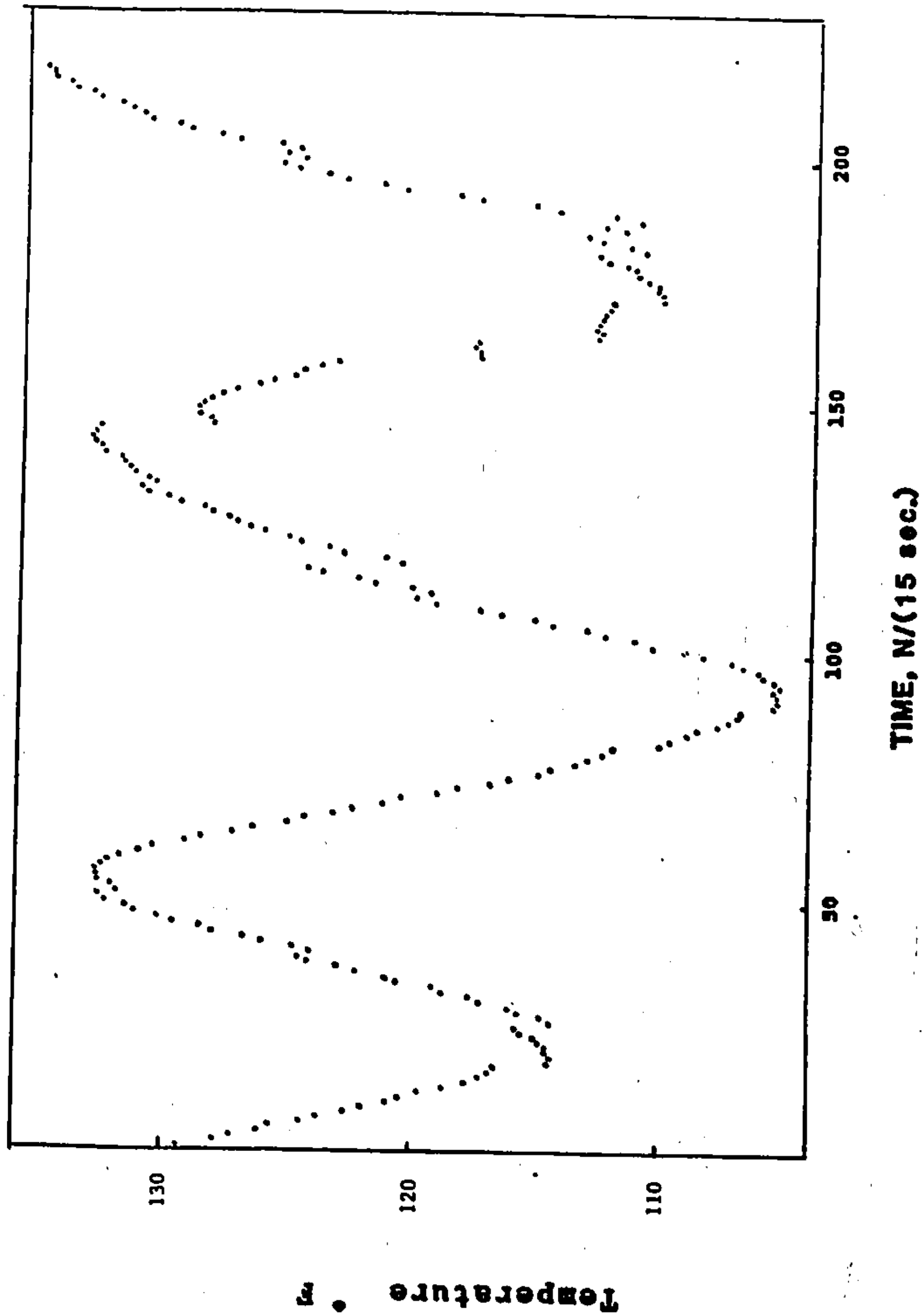


Figure 26: Normal Closed-Loop Operating Data: Outlet Temperature

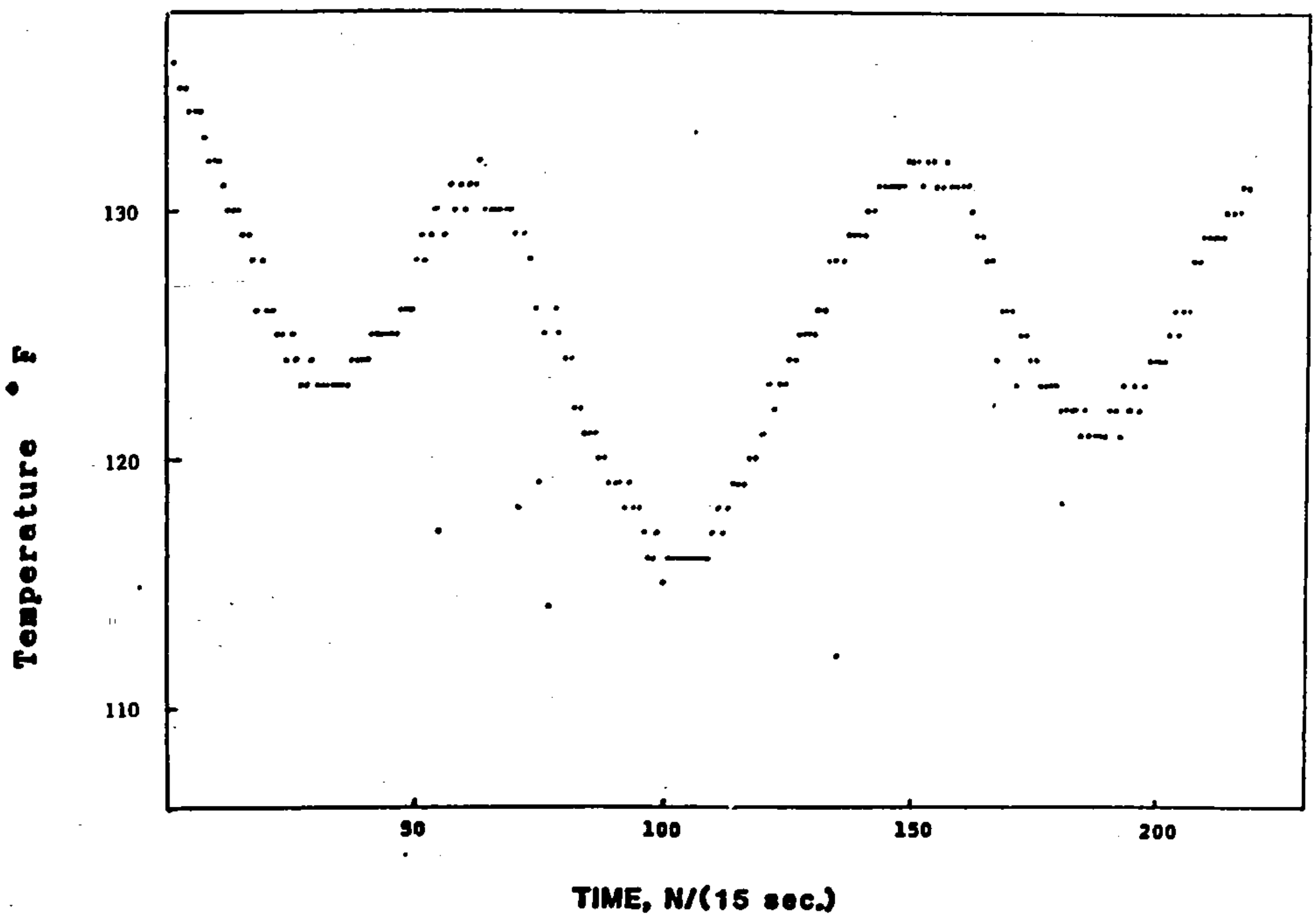


Figure 27: Normal Closed-Loop Operating Data: Drum Temperature

is not consistent. The manipulated variable, for the closed-loop data illustrated in Figure 24, exhibits the same sinusoidal behaviour as observed for the system temperatures, while the manipulated variable's response for the PRBS open-loop data is a PRBS time series.

Review of the available data suggested that the assumption of open-loop operation for the PRBS data was invalid. A standard test for the presence of a feedback is described in Box and Jenkins [1]. An investigation of crosscorrelation estimates at negative lags will indicate the presence of feedback control. If any of the calculated crosscorrelation estimates at negative lags are significantly different from zero then the data were observed under closed-loop operation with respect to the measured manipulated variable. No apparent feedback with respect to the input was detected for the PRBS data.

A comparison of the PRBS and closed-loop temperature responses suggests strongly the presence of feedback control for the PRBS data. However, a diagnostic check of the PRBS data for feedback control does not substantiate this argument. This comparison also suggests a second argument: that the system input provided is not the actual input of the system but only produced in the software. Ignoring the measured PRBS input signal, the measured system temperature responses could quite easily be mistaken for data observed under closed-loop operation. The assumption of inadequate input data would invalidate the diagnostic check for feedback control since the actual system input is unknown.

The evidence of the available data, both under closed-loop and open-loop conditions, and the possibility of feedback control strongly supported the hypothesis of inadequate data collection as a probable cause for the inconsistencies in the results. Box and MacGregor [14] state that if the assumption of open-loop operation is not valid, system identification will be compromised.

6.2 Variation of Process Model Order / Parameters Between Data Collection Periods

System data, collected at different time periods, will produce model forms and parameter estimates of the true system which are not identical. Variation between the calculated models is expected and considered normal for most chemical processes. However these variations are expected to be small. Drastic modification of the system parameters requires the real modification of the process. The airfin heat exchanger system did not undergo major maintenance or modifications. Therefore the inconsistencies in the observed results cannot be explained by variations in the system order or parameters between data collection periods.

6.3 Verification of Algorithms

All relevant calculations for the Box Jenkins model building technique were performed using user-written software. An obvious cause of inconsistency may have been inaccuracies in the software.

The algorithms consisted of three distinct and separate programs, an autocorrelation function estimation routine, a crosscorrelation function estimation routine, and a grid search routine. All the routines were tested and modified until deemed acceptable by the user. The iterative approach, as outlined by Box-Jenkins was performed using these routines. All interpretations, and decisions of the iterative approach were performed off-line, the algorithms providing only the necessary mathematical calculations.

The PRBS open loop data investigation required only the use of the autocorrelation and crosscorrelation function estimation routines. The results, the calculations and routines used to produce them were analyzed and double checked several times before the conclusions of the initial investigation were presented to Imperial Oil. No errors were found. Additional support for the validity of the routines was provided by comparison with results obtained using the time series analysis algorithms developed at McMaster University. The results of the second analysis agreed with the original results.

A qualitative graphical interpretation of the system using the step response data indicated that each of the three temperature responses to the step function input could adequately be described by a first or second order system with dead time. Analysis of the step response data using the Box-Jenkins technique and the user developed software yielded similar results. The proposed models in Table 2 for each of the three bivariate stochastic processes are analagous to a second order plus dead time system representation. Therefore the observed inconsistencies of the results could not be adequately explained or accounted for by program error in the algorithms.

6.4 Inadequacies of the Box-Jenkins Technique

The identification stage of the Box-Jenkins procedure produced the major inconsistency in observed results. Statistical independence between the manipulated variable and both the heat exchanger outlet stream temperature and drum temperature was indicated using the PRBS data. On the other hand, the same identification procedure when applied to the step response data identified a significant dependence between the manipulated variable and the two temperatures.

It was then assumed that significant dependence did in fact exist. This was consistent with the physical reality of the system and strongly contended by Imperial Oil. On this basis, the Box-Jenkins procedure failed to recognize the presence of correlation using the PRBS data. The following hypotheses were postulated to explain this failure:

(i) The outlet stream and drum temperature responses were masked or buried in the associated system noise.

(ii) The control interval of 15 seconds was too short and did not allow enough time between perturbations for the system to respond.

A computer simulation program was then developed to investigate the respective effects of system noise and control interval on the identification stage of the Box-Jenkins procedure. The model relating the outlet stream temperature to the manipulated variable, determined from the step response analysis, was used as the system transfer function for the computer simulation.

The simulated system was perturbed, under open loop conditions, with a PRBS input sequence of equal amplitude to the PRBS input of the observed data. Noise, of known variance, was also added into the simulated system. The variance of the added noise ranged from zero to more than double the calculated variance of both the input PRBS perturbation series and step function input. The sampling rate and control interval of the PRBS input series were set at values of 15 seconds, 30 seconds and 1 minute. The sampling rate was always equal to or less than the control interval. The transfer function model derived with a sampling rate of 1 minute was transformed into an equivalent representation for sampling rates of 15 and 30 seconds using the procedure outline by MacGregor [22].

The time span for the simulation was held constant. Therefore, the number of samples used in the identification procedure increased with shorter sampling times. Approximately one hour of data was simulated yielding a maximum of 240 simulated observations. This number was consistent with the number of samples used in the PRBS data analysis which varied between 225 and 240.

Results of the simulation study are tabulated in Table 3. The identification step of the Box Jenkins procedure recognized significant correlation between the input and output variables even when the noise variance was ten times greater than that of the simulated system response. When identifying system models, information contained within the observed data may be lost due to the presence of noise. System noise will mask the actual system response thereby producing a distorted view of the system. Even though the transfer function identified under very noisy conditions may not accurately represent the true system transfer function, the simulation study demonstrated clearly that at least the existence of a transfer function could be identified at very high noise levels. The results of the simulation investigation support the rejection of the first hypothesis as a possible explanation for the observed inconsistencies.

Table 3: Simulation Results for the Effect of Sampling Rate, Control Rate and Noise Level on The Box-Jenkins Technique's Preliminary Identification

Pulse Rate of System Input Sampling Interval	Variance of Noise added to simulation Data						
	1.0	4.1	9.3	16.6	25.9	37.4	50.9
Pulse rate: 1 min Sampl. rate: 1 min	+**	+	+	+	+	+	+
Pulse rate: 1 min Sampl. rate: 30 sec	+	+	+	+	+	+	+
Pulse rate: 30 sec Sampl. rate: 30 sec	+	+	+	+	+	+	+
Pulse rate: 30 sec Sampl. rate: 15 sec	+	+	+	+	+	+	+
Pulse rate: 15 sec Sampl. rate: 15 sec	+	+	+	+	+	+	+

Please Note: Variance of model = 19.77 for 1 minute sampling interval
Variance of model = 6.9 for 30 second sampling interval
Variance of model = 3.69 for 15 second sampling interval

** The sign (+) indicates that a transfer function model was detected with the Box-Jenkins preliminary identification procedure.

The identification stage detected significant correlation between the system input and output variables for all values of the control interval. Significant system responses to changes in the manipulated variable were observed at the fastest control rate of 15 seconds. The simulation study indicated that the control interval of 15 seconds allowed sufficient time for the system to respond and therefore some correlation or transfer function should have been evident using the PRBS data. Thus this cause appears unlikely in explaining the differences in results obtained from the two analyses.

6.5 Conclusions

Possible causes to explain the contradiction in observed results were proposed and investigated. Possible errors in the computer algorithms developed for estimation of the autocorrelation and crosscorrelation functions were considered and rejected. The algorithms' results were verified independently using the same PRBS open loop data and identical conclusions were derived. Furthermore the user developed algorithms successfully identified system models for the step function open loop data.

The inability of the Box-Jenkins technique to identify system dependence in the presence of noise was examined. Simulation studies concluded that the Box-Jenkins approach was capable of identifying correlation between system variables in the presence of high noise levels and at the preset sampling and control intervals. It was concluded that this explanation was not plausible.

Significant variations in system model order/parameters between data collection periods could account for the inconsistencies. However, actual physical changes to the real system would be required. No major modifications were performed on the system. Furthermore, consistent results between the normal open loop and step function open loop data sets (separated in time by more then six months) were presented. Therefore this cause for the discrepancies observed was rejected also.

The possibility of inadequate data was postulated. Review of the available data sets support this allegation. Consistent results were observed for 3 of the 4 data sets. The PRBS data, which produced the contested results, was the lone exception. The behaviour of the PRBS data indicated inadequate input data and the possible violation of the stated condition of open loop operation. Therefore the presence of inadequate data, specifically the PRBS data set, is accepted as the probable cause of the observed discrepancies in the analyses.

7.0 CONCLUSIONS AND RECOMMENDATIONS

The Box-Jenkins technique was applied to an industrial system identification problem. Plant data from an airfin heat exchanger system was provided by Imperial Oil. The major objective of the thesis was the investigation of the applicability and the effectiveness of the Box-Jenkins approach for system identification in the industrial milieu. The following conclusions can be drawn from this investigation:

(i) The Box-Jenkins model building technique, though requiring some experience for the interpretation of the identification stage, is a simple iterative approach which is easily implemented.

(ii) The Box-Jenkins technique is appropriate for industrial applications. Simulation studies showed that system identification can be achieved in the presence of significant system noise.

(iii) The manipulated variable, percentage of operating pressure to the louvres, is a reasonable choice for the control variable and a stable control strategy is therefore feasible.

(iv) The inconsistency of the results between the PRBS and step response data sets appears to have been the result of feedback control actually existing when the PRBS data were collected.

The Box-Jenkins analysis of the PRBS data indicated no significant relationship between the manipulated variable and the outlet and drum temperatures. These results were inconsistent and unexpected with respect to the physical reality of the system.

A second system identification analysis was performed using step response data provided by Imperial Oil following the presentation of the PRBS data results. The identical step by step Box-Jenkins analysis was applied to the the new data. The step response analysis indicated strong correlation between the manipulated variable and the three temperature variables. This was the expected system behaviour. However, meaningful conclusions concerning the airfin heat exchanger system could not be made without an explanation of the inconsistencies in observed results.

A review of possible causes for the contradictory results was initiated. The probable cause was concluded to be the PRBS data. The PRBS data did not provide an adequate system description under the stated condition of open-loop operation.

Considerable time and effort were expended to determine the presence of unreliable data. Furthermore, the confidence in all of the data provided by Imperial Oil was undermined by this finding. Verification of the inadequqacy of the PRBS data provided was not feasible for this investigation since the system in question, the airfin heat exchanger system, has been shut down indefinitely by the Refinery.

The effectiveness and applicability of the Box-Jenkins technique for industrial use cannot be adequately judged using step response data. Without reliable data and the impossibility of additional data collection for analysis, development of a computer simulation of the system and subsequent adequate control strategy using the Box-Jenkins technique could not be performed. However, considerable insight into the practical aspects of the Box-Jenkins technique and its effectiveness as a system identification and control tool were gained by this investigation. Furthermore, the difficulties of interfacing with the industrial milieu were recognized and the condition of significant input or control of the data collection stage is now considered a prerequisite for any collaborations in the future.

The following recommendations are proposed:

(i) If possible, the further substantiation of the presence of inadequate data should be undertaken with new measured data from the system under study.

(ii) Great care should be exercised during the data collection phase in the industrial environment so that the data accurately reflects the system under the conditions stated.

8.0 REFERENCES

[1] Box, G.E.P. and Jenkins, G. "TIME SERIES ANALYSIS: FORECASTING AND CONTROL", 2nd, Holden Day, San Francisco, 1976

[2] Eykhoff, P. "SYSTEM IDENTIFICATION, PARAMETER AND STATE ESTIMATION", Wiley, 1974

[3] Hsia, T.C. "SYSTEM IDENTIFICATION", Lexington, 1977

[4] Isermann, R. "New Results on the Identification of Processes", Automatica,7, p. 191, 1971

[5] Isermann, R. "Practical Aspects of Process Identification" Automatica, 7, p. 595, 1971

[6] Astrom, K.J. and Eykhoff, P. "System Identification - A Survey" Automatica,7, p. 123, 1971

[7] Graupe, D. "IDENTIFICATION OF SYSTEMS" 2nd, Krieger, Florida, 1976

[8] MacGregor, J.F. "Process Identification" Volume 2 of "Advanced Process Control, An Intensive Short Course on Digital Computer Techniques for Process Identification and Control" McMaster University, 1982

[9] Schwartz, R.J. and Friedland, B. "LINEAR SYSTEMS" McGraw Hill, 1965

[10] Nyquist, H. "Regression Theory" Bell System Journal, 11, p. 126, 1932

[11] Bode, H.W. "NETWORK ANALYSIS AND FEEDBACK AMPLIFIER DESIGN" Van Nostrand Reinhold, New York, 1945

[12] Clarke, D.W. "Generalized Least Squares Estimation of the Parameters of a Dynamic Model" IFAC Symp. Ident. in Aut. Contr. Sys., Prague, paper 3.17, 1967

[13] Isermann, R., Baur, U., Bamberger, W., Kneppo, P. and Siebeit, H. "Comparison of Six On-Line Identification and Parameter Estimation Methods" Automatica, 10, p. 81, 1974

[14] Astrom, K.J., Bohlin, T. and Wensmark, S. "Automatic Construction of Linear Stochastic Dynamic Models for Industrial Processes with Random Disturbances using Operating Records" Report IBM Nordic Lab, Stockholm (1965)

[15] Bohlin, T. "On the Problem of Ambiguities in Maximum Likelihood Identification" Automatica, 7, 1979

[16] Saridis, G.N. "Comparison of Six On-Line Identification Algorithms" Automatica, 10, 1974

[17] Saridis, G.N. and Stein, G. "Stochastic Approximation Algorithms for Linear System Identification" IEEE Trans. Auto. Contr., AC-13, 1968

[18] Chatfield, C. "THE ANALYSIS OF TIME SERIES: AN INTRODUCTION" 2nd, Chapman and Hall, 1980

[19] Astrom, K.J. and Bohlin, T. "Numerical Identification of Linear Dynamic Systems from Normal Operating Records" from "THEORY OF SELF-ADAPTIVE CONTROL SYSTEMS", Hammond, P.M. Ed., Plenum Press, New York, 1977

[20] Goodwin and Payne "Dynamic System Identification: Exp. Design and Data Analysis" Math in Sci. and Eng., 136, Academic Press, 1977

[21] Anderson, O.D. "On the Collection of Time Series Data" Opl. Res. Q. 26, p. 331, Pergamon Press, Great Britain, 1975

[22] MacGregor, J. F. "Optimal Choice of the Sampling Interval for Discrete Process Control" Technometrics, 18, p. 151, 1976

[23] Bacon, D.W. and Howe, G.L. "Statistical Evaluation and Design of an Evaporator Control System" Can. Journal of Chem. Eng., 50, 1972

[24] Box, G.E.P. and MacGregor, J.F. "The Analysis of Closed-Loop Dynamic-Stochastic Systems" Technometrics 16, p. 391, 1974

[25] Phadke, M.S. and Wu, S.M. "Identification of Multi-input-Multi-output Transfer Function and Noise Model of a Blast Furnace from Closed-Loop Data" IEEE Trans. on Auto. Contr., AC-19, p. 944, 1974

[26] Wright, J.D. and Bacon, D.W. "Analysis of a Heat Exchanger Network Using Statistical Time Series Analysis Methods" Ind. Eng. Chem., Process Des. Dev., 14, p. 453, 1975

[27] Goford, P., MacGregor, J.F. and Wright, J.D. "Predictive Stochastic Feedforward-Feedback Control of a Heat Exchanger-stirred tank system" Int. J. Contr., 26, p. 463, 1977

APPENDIX 1

ADDENDIX 1

ADAPTIVE CONTROL

Adaptive control is not a new concept, it was first proposed more than 25 years ago. Rather, it is an old idea which has received considerable attention during the last decade. From its conception in the late fifties until the early seventies several adaptive control schemes were proposed. However, progress in the field was relatively slow and adaptive control remained a province of experience and art. Pratical applications were scarce during this time period and no consistent theory of what constituted an adaptive control system was forthcoming. The last decade has produced the consolidation of sound theories for adaptive control. Major breakthroughs include the resolution of the long standing stability problem and the realization of the equivalence of various approaches for adaptive control.

Progress in the development of adaptive controllers was also retarded by the cost and complexity of implementation of such schemes. The advent of microprocessors has provided both the necessary reduction in cost and ease of implementation required.

With the resolution of the stability problem and available computer power, adaptive control is now an acceptable alternative for resolving system control problems in industrial applications.

Model Reference Adaptive Control (MRAC) and Self-Tuning Regulators (STR) are the principle approaches in the adaptive control literature. STR systems are also referred to as parameter adaptive systems. In the MRAC approach, the objective is to make the output of an unknown plant asymptotically approach that of a given reference model. MRAC systems are less complex in their structure then STR systems, they include only one computer procedure, namely the minimization of the error between the output of the process and the reference model. In the STR technique, a design procedure for known plant parameters is first determined and is then applied to the unknown plant using recursively estimated values of these parameters. The STR approach is therefore a two-step procedure, namely identification and optimization. Similarities between the two techniques have been noted. Study of these similarities in recent years have demonstrated the equivalence of the two techniques when the same criteria for the adaptive system are used.

The objective of this survey is to provide an overview of adaptive control and its present status of development. The adaptive control problem is discussed in section 1. The two major approaches, STR and MRAC, are reviewed in sections 2 and 3. Similarities and the equivalence of the two techniques are presented in section 4. Stability and convergence of adaptive controllers are treated briefly in section 5. Section 6 reviews adaptive control applications published in the literature. A discussion of unresolved questions of both theoretical and practical interest is presented in section 7.

A1.1 The Adaptive Control Problem

Adaptive controllers have evolved as a technique to avoid degradation of the dynamic performance of a control system when environmental variations occur. The control engineer's objective, in practice, is the design of the simplest control scheme which will achieve the desired control criteria in spite of the variations of the dynamic parameters. In the presence of small parameter variations, a fixed control law is generally adequate. When acceptable performance cannot be achieved due to large variations in the environment and/or incomplete knowledge of the system in question, an alternate approach is required.

A new class of control schemes refered to as adaptive control have been developed to provide system control in the presence of large and unpredictable variations in system parameters and/or incomplete knowledge of th system in question. The adaptive control scheme constitutes a branch of optimization theory dealing with a specific type of control problem: design of an optimal control scheme with incomplete system model knowledge in real time where conventional fixed parameter controllers prove inadequate.

Definition of Adaptive Control

Separation of control schemes into either adaptive control schemes or non-adaptive schemes is not clear cut. No single definition of adaptive control, of those presented in the literature, has received general acceptance. Definitions vary from a very broad general definition to a specific definition of the control scheme. The following definitions have been drawn from the literature and are representative of the wide scope of defintions.

i) "A adaptive system is any physical system which has been designed with an adaptive viewpoint" [A1]

ii) "...an adaptive system would automatically compensate for variations in system dynamics by adjusting the controller characteristics so that the overall system performance would be satisfactory. Such a system would include elements to measure or estimate the process dynamics, and other elements to change the controller characteristics accordingly" [A2]

iii) "...control in which automatic and continual measurement of the process to be controlled is used as a basis for the automatic and continuing self design of the control system" [A1]

iv) "An adaptive system measures a certain index of performance (IP) using the inputs, the states, and the outputs of the adjustable system. From the comparison of the measured index of performance and a set of given ones, the adaptation mechanism modifies the parameters of the adjustable system or generates an auxiliary input in order to maintain the index of performance close to the set of given ones" [A3]

It is not the purpose of this work to select any one definition. The importance, as supported by the literature, of the adaptive control concept is that it helps us to consider a whole new area of challenging problems. A good indication of adaptive control action, when examining control schemes, is the ability of the control scheme to cope with an environment which a time-invariant system could not handle.

The Adaptive Control System

If knowledge of the process model is incomplete because of random time-varying changes in the environment or of the parameters, then the initial identification, decision and modification will not be sufficient to optimize the performance index. Optimization of the system will require repetition of the procedure continuously or at set time intervals. This constant reevaluation of the system is necessary to counteract unpredictable changes in the process. The ability to modify the control strategy is usually considered in defining an adaptive control system.

The adaptive control process may be broken down into three major functions: identification, decision and modification. Although it may be difficult to separate those parts of the system responsible for each, all three are required for adaptive action to take place. An adaptive controller is then an algorithm consisting, in general, of the following functions:

1) Identification of dynamic characteristics of the plant or process
2) Decision of modification required
3) Implementation of required modification

Some authors describe the essential components of adaptive control by only two elements: identification or estimation and actuation or control law, the last element incorporating the decision and modification stages. But in recent years the former breakdown of the adaptive control system has become prevalent.

Adaptive Control and Self Tuning Regulators - Their Equivalence

When discussing adaptive control, a term commonly used for describing a similar or identical system is the self tuning regulator (STR). Although most authors agree now that the two terms are interchangeable, STR systems were originally designed for the resolution of a simpler control problem. Astrom and Wittenmark [A4] state the following:

> "For systems with constant but unknown parameters it thus seems reasonable to look for strategies that will converge to the optimal strategies that could be derived if the system characteristics were known. Such algorithms will be called "self-tuning" or "self-adjusting" strategies. The word adaptive is not used since adaptive...usually implies that the characteristics of the process are changing."

But Astrom expands the applicability of STR further in a subsequent article [A5] to include systems with slowly varying parameters.

It is therefore not necessary to assume that the parameters are constant and that we can generalize that the parameters of the system are stochastic processes without increasing the complexity of the problem. STR is accepted presently as one of the most popular approaches to adaptive control, the other being model reference adaptive control (MRAC). The similarities of the two approaches are discussed in chapter 4.

A1.2 Self-Tuning Regulators

Stochastic control theory has been proven to be a very useful and practical tool for the design of controllers for industrial processes. However, there exists many practical applications where it is difficult to determine the parameters of the controller since the dynamics of the process and its disturbances are unknown. In such cases it is then necessary to go through the steps of plant experiments, parameter estimation, computation of control strategies and implementation. The experiments and their subsequent evaluation can be time consuming which is undesirable. Furthermore, repetition of the experiments may be required if the plant parameters or disturbances are changing with time. It is thus desirable to have a regulator which tunes its parameters on-line. The self-tuning regulator can be regarded as a convenient way to combine system identification and control design. STR control strategies negate the need for plant experimentation. Furthermore, if changes in the process parameters or disturbances are not too rapid, a STR controller will provide close to optimal control of the system.

The STR controllers are based on a fairly natural combination of identification and control. The parameters of the system controller are unknown. They are therefore obtained from a recursive parameter estimator. A separation between identification and control is assumed. The only information used from the estimator by the control law are the parameter estimates. The fact that the parameter estimates are not exact is disregarded. This implies that a certainty equivalence control is being implemented. There exists control strategies which take into account the uncertainties of the estimates but these strategies are not considered here.

The general configuration of a STR system is shown in figure A1.2.1. The regulator can be visualized as being composed of three parts, a recursive parameter estimator, a design calculator and a controller or regulator with adjustable parameters. The design calculator relates the controller parameters to the parameter estimates which describe the process. The parameter estimates characterize the process and its environment from the measured process input and output. The configuration in A1.2.1 is for a single-input single-output system. The discussion of this

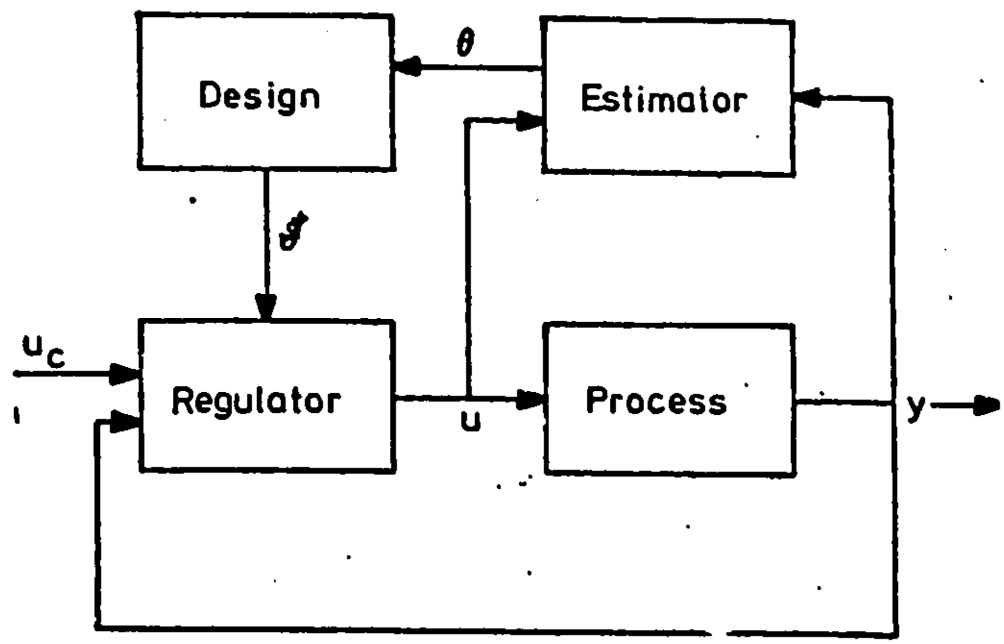


Figure A1.2.1: STR Structure

appendix will be limited to the single-input single-output case.

Probably the first to formulate this simple idea as an algorithm was Kalman [A6]. An on-line least-squares routine calculated estimates of plant parameters. The control law was then calculated at every sampling interval using the estimates of the least-squares routine.

The STR concept was developed by Peterka [A7] and Astrom and Wittenmark [A4] in a stochastic framework. Astrom and Wittenmark's algorithm, based on minimum variance control, is presented below as an example of a STR.

Consider the following plant

$$y(t) = ay(t-1) = bu(t-1) + e(t) \qquad \text{(A1.1)}$$

where u = input
y = output
$\{e(t)\}$ = sequence of independent, zero-mean random variables

For this plant, the following control law will give the minimum output variance

$$u(t) = (a/b)y(t) \qquad \text{(A1.2)}$$

If the parameters a and b are not known, the algorithm developed by Astrom and Wittenmark can be applied. This algorithm consists of two steps, which are both repeated at every sampling instant:

1) Estimation of the parameters α and β_0 in the model

$$y(t) = \hat{\alpha}y(t) + \hat{\beta}_0 u(t-1) + \epsilon(t) \qquad \text{(A1.3)}$$

2) Implementation of the control law

$$u(t) = -[\hat{\alpha}(t)/\hat{\beta}_0]\, y(t) \qquad \text{(A1.4)}$$

where $\hat{\alpha}$ and $\hat{\beta}_0$ are the estimated values of α and β_0. It should be noted that the estimation of the parameters can be made recursively if a least-squares criterion is used. This makes the scheme practically feasible. The final conclusion of the analysis of this system by Astrom and Wittenmark [A4] and Ljung [A8] is that the algorithm will converge under a stability condition if the noise characteristics satisfy a certain positive realness condition. Goodwin et al [A9] obtain similar results without the stability assumption.

The general structure of the STR encompasses a variety of schemes since there exists many different ways to do control and estimation. The STR approach is therefore not limited to the minimum variance control strategy. Other approaches have been proposed. Astrom and Wittenmark [A10] and Clark and Gawthrop [A11] proposed generalizations of the basic algorithm. Algorithms based on pole placement design were discussed by Edmunds [A12], Wellstead et al [A13] and Astrom et al [A5]. Astrom [A15] concludes that although there exists a large variety of possible combinations of design methods and parameter estimation algorithms, only a small number of those available have been explored.

Explicit and Implicit Self-Tuning Regulators

Algorithms where the plant parameters are estimated and the control parameters adjusted on the basis of these estimates are referred to as explicit self-tuning regulators.

The design calculations required for the explicit algorithms may be time consuming. Algorithms have been developed which rewrite the process model such that the design step becomes trivial. Therefore the design calculations are avoided and the parameters of the regulator are updated directly with the proper choice of model structure. This type of algorithm is referred to as implicit self-tuning regulators.

Conclusions

A systematic approach to the design of STR has been presented. The basic concept is very straightforward: a design procedure for a system with known parameters which fits the particular application is initially chosen. If the parameters of the plant are not known, they are estimated recursively and the control parameters are calculated at each interval using the updated estimates. Proper choice of plant model can eliminate the design calculation step.

There are several theoretical problems associated with the STR approach and adaptive control in general. Stability, convergence and performance are the major concerns and are discussed in subsequent chapters.

A1.3 Model Reference Adaptive Control

Adaptive control schemes were introduced to maintain acceptable performance in the presence of environmental variations. While a feedback control system is oriented toward the elimination of the effect of state perturbations, the adaptive control system is oriented toward the elimination of the effect of structural perturbations upon the control strategy. These structural perturbations are caused by changes in the dynamic parameters of the controlled plant. In this section the MRAC approach to adaptive control will be reviewed.

The area of model reference adaptive control is more difficult to characterize in a general way. The difficulty is that the many different schemes proposed were motivated by different considerations. However, a basic configuration can be described. In MRAS (model reference adaptive systems) the specifications are given in terms of a reference model which tells how the process output should respond to the command signal. The unknown plant is controlled by an adjustable controller. A schematic diagram of a MRAS is presented in figure A1.3.1. The reference model is shown as part of the control system in figure A1.3.1. The controller can be considered to have two loops: the inner loop is an ordinary control loop and the outer loop modifies the parameters of the regulator in the inner loop by some kind of adaptation mechanism in such a way that

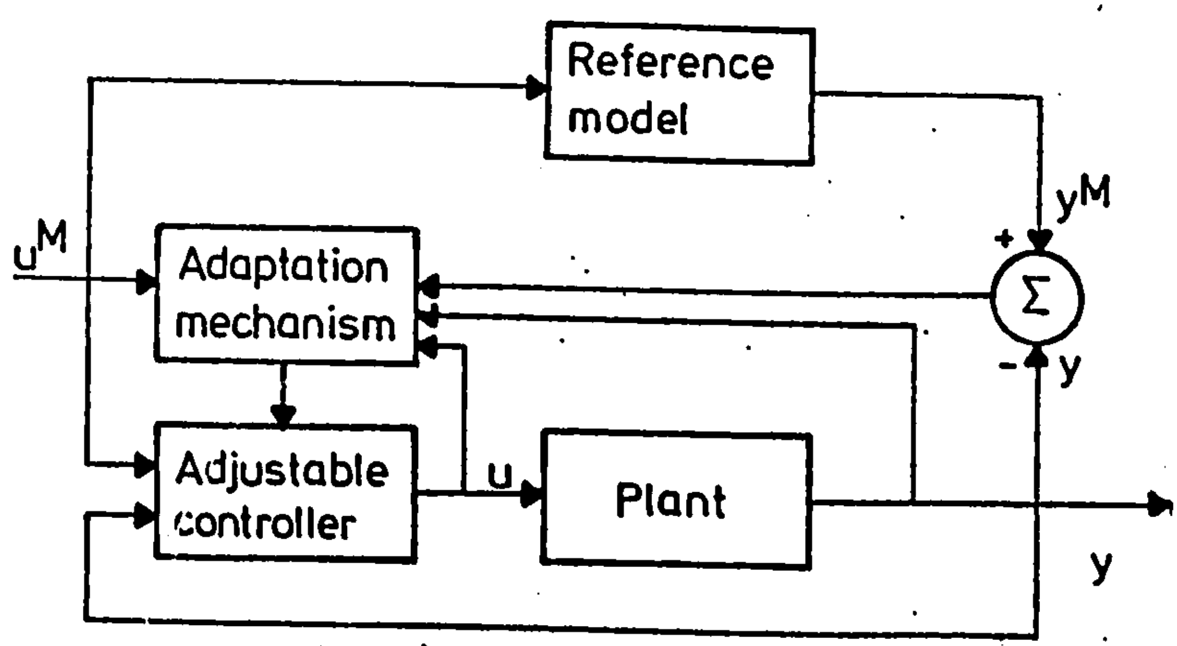


Figure A1.3.1: MRAS Configuration

the error between the process output and the model output is minimized.

One should note that the implementation of the three fundamental blocks is complex and a distinct separation of the adaptive system according to figure A1.3.1 is not always straightforward. However, the presence of a closed-loop control on the identification block is indicative of the presence of adaptive control. Control schemes using open-loop identification and identified as adaptive control are discussed in more detail later in this section.

<u>Indirect and Direct MRAC</u>

Two philosophically different approaches exist for the solution of the MRAC problem. In the first approach, indirect MRAC, the plant parameters are estimated and the control parameters are modified based on these estimates so that the overall plant transfer function matches the reference models transfer function. The identification of plant parameters is eliminated in direct control and the control parameters are directly adjusted to minimize the error between plant and model output. Indirect MRAC control is a form of STR in which the objective is to reduce the error between plant and model outputs.

<u>Open-Loop Model Reference Adaptive System</u>

In the open-loop structure of adaptive control, the decision process is reduced to a fixed mapping ot the process parameters to the controller parameters. The original decision process is already realized in the design phase of the adaptive control system. In these cases, where knowledge of the system is present, it is therefore possible to design preprogrammed time variations of controller parameters to achieve instantaneous optimum control at all times. This type of adaptive control system assumes a rigid relationship betwen some measurable variable of the environment and the dynamic parameters of the system. It is referred to as open-loop adaptive control because the modifications on system performance operated by means of the adaptive mechanism are not measured and fedback to the comparison-decision block. The approach is also called preprogrammed adaptive control. A typical application of this approach is the modification of the characteristics of an autopilot of an airplane using speed and altitude measurements for each flight configuration. In reality, of course, such a controller is simply an optimum time-varying system, the control parameters being automatically changed to their optimum values relative to instantaneous environmental conditions encountered at each point in time.

Classification of Model Reference Adaptive Systems

There exist many MRAC configurations, which cover a broad spectrum of possible applications. It is therefore impossible to consider only one criterion when trying to establish the classification of all the typical structures. But, by considering several critieria a given configuration can be related to some typical MRAS. Landau [A3] offers the following possibilities for specific classification:

1) structure
2) index of performance
3) type of application
4) type of parameter disturbance

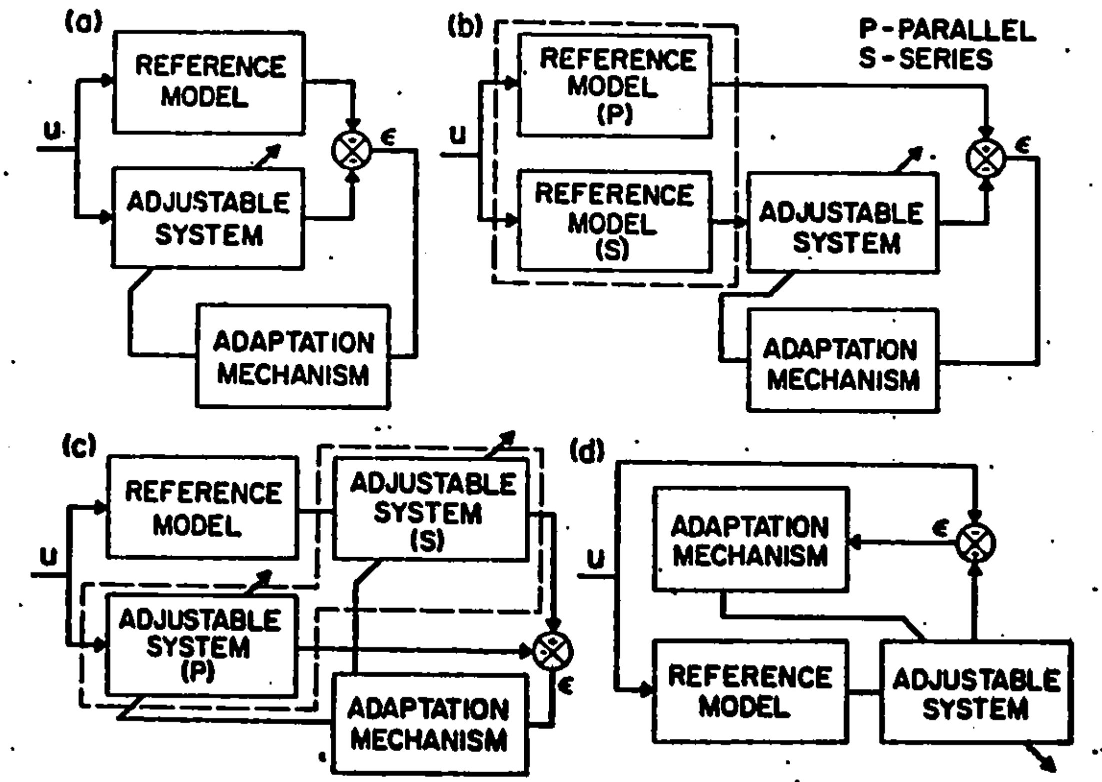
Structure refers to the placement of the reference model within the adaptive control loop. Three basic structures are presented by Landau [A3] and can be found in figure A1.3.2.

Similar to the STR approach, where the combination of identification methods and control laws provided a enormous array of possible algorithms, so too the combination of the last three critiria, index of performance, type of application and type of parameter disturbances, which create a large range of different syntheses of MRAS schemes.

Although a variety of MRAC schemes exist in the literature, the presence of a reference model, an adjustable system and an adaptation mechanism is required for any proposed scheme to be referred to as a MRAC system.

Conclusions

The basic MRAS control scheme has been described. The presence of a reference model, an adaptation mechanism and



Basic structures of MRAS's. (a) Parallel. (b) and (c) Series-parallel. (d) Series.

Figure A1.3.2: Reference Model Structure an adjustable controller are required for consideration as a model reference adaptive system. The distinction between a closed-loop MRAC system and an open-loop MRAC system is the absence of feedback to the adaptation mechanism of the effects provocked by controller manipulation. Finally, the immense diversity of MRAC systems is shown. Relevant theoretical and pratical problems are presented in the following chapters.

A1.4 Equivalence of the STR and MRAC approaches

The MRAC and STR approaches represent two broad areas of adaptive control. The STR and MRAC were originally developed to solve different control problems. The STR was conceived to solve the stochastic regulator problem. The MRAC was designed to solve the deterministic problem.

In MRAC the first step is the choice of a reference model followed by the selection of an approximate controller structure. For the STR, a design procedure is initially determined which can be used when the plant parameters are known and this is applied to the unknown plant using recursively estimated values of the parameters. The tendency in MRAC to apply stability theory in the design procedure is not present for the STR approach. While MRAC has been applied principally in continuous time systems, recently several discrete time applications have been initiated. The STR has generally been analyzed and implemented only in discrete time. In spite of these differences the two techniques exhibit some important similarities.

Many authors have suggested that STR and MRAC have many common characteristics. Recently, researchers in both disciplines have become more aware of the interrelation between them. The similarities have been examined more precisely and efforts have been initiated to show the equivalence of the two schemes. Many of the proposed STR and MRAC schemes presented in the literature have been shown to be the same or special cases of one algorithm. Astrom [A16], Matko [A16] have demonstrated that in systems using the same criterion, both approaches lead to identical regulator structure. The basic algorithm for MRAC and STR is shown to be equivalent in [A17], [A18] and [A19]. Egardt [A18] has presented an unified approach with MRAC and STR being special cases of a fairly general algorithm. However, it is not always shown that the resulting systems are globallyl stable.

In conclusion, all the globally stable adaptive schemes presently known are equivalent in that they have the same error equations and the same adaptive laws. Several independant analyses have demonstrated one common approach.

A1.5 The Stability and Convergence Problems of Adaptive Control Schemes

STR Approach

Two key areas of concern for the STR are the questions of overall stability and convergence. The stability problem can be viewed in several different ways. For example, the study of local stability and its results are presented by Feuer and Morse [A20]. But this approach is of limited practicality and interest since it reveals very little about the global properties. The properties of global stability are inherently more difficult to investigate. One approach to resolve the question is to apply Lyaponuv theory, but this technique suffers from the difficulty of finding an appropriate Lyaponuv function.

In the past few years, considerable progress has been made in stability theory for STR. Unfortunately, most of the results are limited to the simple STR based on least-squares estimation and minimum variance control [A4] or its direct multivariable generalization [A21] [A22]. Stability results have also been proven for simple STR using pole placement. The simple self-tuner will work well only for minimum phase systems. Conditions for stability when there are no disturbances are given in [A9]. The corresponding results for the bounded disturbances' case are found in [A23]. The theoretical results are also limited due to the assumptions made. These assumptions, required by the theory, are difficult to verify in a practical situation. The most restrictive assumption is the presence of an upper bound of the order of the system that must be known. This is unrealistic in practice where the process to be controlled will generally be of high order and the STR will be based on a simplified model. It has been shown, however, that STR based on drastically simplified models of the system actually work very well [A24].

Overall stability of the closed-loop sytem is perhaps the most fundamental property but the problem of convergence is also important. Convergence for simple STR based on least-squares and minimum variance control has been investigated [A25]. The surprising result of the study, that the regulator is the only possible equilibrium point, even in the presence of a structural mismatch between the model and the process, is supported in [A4].

In summary, research in recent years have developed stability and convergence criteria for the STR. However, elimination of some of the required assumptions and extension of the analysis beyond the most simple case is necessary.

The MRAC Approach

Stability problems are inherent in the MRAC approach due to its time-varying non-linear character. Therefore, design of a MRAC system requires a stability analysis. Previous attempts to design MRAC systems have shown the difficulty of examining global stability characteristics when local parametric optimization is used in the adaptation mechanism.

The methods to design the adaptation loop in MRAC has mostly been based on stability theory since the work of Parks [A26]. Parks initiated a rapid development of the MRAC approach by applying stability theory and using Lyaponuv functions in the design stage.

The literature supports the formulation of the MRAC control design as a stability problem and the use of suitable techniques for solution of this non-linear, time-varying stability problem. The control law determined by this approach is very practical since a major priority of any satisfactory MRAC system is the overall stability of the system.

A survey of various MRAC systems using Lyaponuv designs for resolution of the stability problem [A27], [A28] states that this approach introduces the question of how to choose a class of Lyaponuv functions in order to increase the class of adaptation laws which yields globally stable MRAC systems. This is an important concept since a wider choice of adaptation laws will provide a higher probability of finding a suitable control law for any one particular application.

A second approach to MRAC design is the resolution of the stability question using hyperstability theory in conjunction with the properties of the positive dynamic system. A description of this approach is given by Laudau [A3].

Analysis of MRAC systems that are both non-linear and time-varying is quite involved, even for the deterministic control problem. Furthermore, a complete analysis of the stability problem in the presence of noise is significantly more difficult although some research is currently in progress.

In conclusion, while the current literature provides theoretical results which establish a sound basis for further exploration, many questions remain unresolved.

Conclusions

Substantial advances in the area of stability and convergence of both the STR and MRAC approaches to adaptive control have been demonstrated. However significant questions still remain unresolved. The issue of the limitations of adaptive control is discussed further in chapter 7.

A1.6 Applications of Adaptive Control

A detailed review of adaptive control applications is not within the aspirations of this appendix. The goal here is to present an overview of applications which have occured in the process industries.

Parks et al [A16], in their review paper, state that although there are several survey papers on adaptive control, few are concerned with practical applications.

Belanger [A15] states that the STR approach has received the most attention in the process industries and the lion's share of known applications in this area.

Parks et al [A16] list several successful implementations of adaptive control schemes in the process industries. Control of a refrigerant compressor test plant with 3 inputs and 3 outputs with acceptable control behaviour over a wide range of operating conditions is achieved. Another successful application involved the control of a heat exchanger. Belanger reviews several successful STR applications in the process industries. Adequate control of moisture on a paper machine and regulation of an ore crusher are provided using adaptive control where ordinary PID control proved inadequate. A third application involved the design of a four parameter STR for the control of a distillation column.

Belanger, in his survey paper, describes several factors which have retarded the implementation of more adaptive control schemes in the process industries when compared to other areas such as electromechanical and electrical power systems. Generally speaking, a good model using system theory is available for electrical and mechanical systems whereas such is not the case for process control. Furthermore the principal perturbations for process control are disturbances which are generally difficult to measure. Another drawback for the process industries is the inability to measure directly most process variables of interest for control. Finally the need for good control is essential for many electrical systems, such as aircrafts or space vehicules, but most processes, due to their long history, can get along with nothing more than good process operators. However, changes in the required tolerances for most of today's processes now make adequate control mandatory.

The lack of a substantial number of applications is generally true in all areas. Lack of a general adaptive control approach and underlying doubts about the available methods have prevented the use of adaptive control by industrial engineers. Until adaptive control can be shown to be superior, the available time-proven methods will always take precedence.

A1.7 The Future of Adaptive Control

The general consensus of the literature is that adaptive control is now ready for practical application. That is not to say that widespread implementation is recommended. Considerable research is still necessary before routine implementation occurs. Astrom [A16] states that the required theory to implement STR confidently is still unavailable. Elimination of some of the required assumptions and extension of the theory beyond the simplest of STR cases is required. The conclusion that the initial assumptions required are too restrictive is also supported by Peterson [A16]. Parks et al [A16] state that more interaction between industry and research institutes is necessary to expand the adaptive control theory to process control in the real world.

A1.8 References

[A1] Miskin & Braun, "Adaptive Control System" published by McGraw-Hill, 1961.

[A2] Smith, C.L., "Digital Computer Process Control", published by Intext Educational Publishers, 1973.

[A3] Landau, Y.D., "Adaptive Control, The Model Reference Approach", Control and System Theory, Vol. 8, published by Marcel Dekker Inc., 1979.

[A4] Astrom, K.J. & Wittenmark, B., "On Self Tuning Regulators", Automatica, Vol 9, pp 185-199 (1973).

[A5] Astrom, Borisson, Ljung & Wittenmark, "Theory and Applications of Self Tuning Regulators", Automatica, Vol 13, pp 457-476 (1977).

[A6] Kalman, R.E., "Design of a Self-Optimizing Control System", Trans. ASME 80, pp 468-478 (1958).

[A7] Peterka, V., "Adaptive Digital Regulation of Noisy Systems", 2nd IFAC Symposium on Identification and System Parameter Estimation, Prague, Czechoslovakia, 1970.

[A8] Ljung, L., "On Positive Real Transfer Functions and the Convergence of Some Recursive Schemes", IEEE Trans Antom Control AC-22, pp 539-550 (1977).

[A9] Goodwin, Ramadge, Caines, "Discrete Time Stochastic Adaptive Control", Div. of Applied Sciences, Harvard University (1978).

[A10] Astrom & Wittenmark, "Analysis of a Self Tuning Regulator for Nonminimum Phase Systems", IFAC Symposium on Stochastic Control, Budapest, Hungary, pp 168-173 (1974).

[A11] Clark & Gauthrop, "Self-Tuning Controller", Proc. IEEE 122, pp 929-934 (1975).

[A12] Edmunds, "Digital Adaptive Pole Shifting Regulators", PhD dissertation, Control Systems Centre, University of Manchester, England (1976).

[A13] Wellstead, Piager, Zanker, Edmunds, "Self Tuning Pole/Zero Assignment Regulators", Report 404, Control Sys. Centre, University of Manchester, England (1978).

[A14] Astrom, Westerberg & Wittenmark, "Self-Tuning Controller Based on Pole-Placement Design", CODEN : LUTFD2/(TFRT-3148)/1-52/(1978).

[A15] Narendra, K.S. & Monopoli, R.V., Editors, "Applications of Adaptive Control", published by Academic Press, Inc., 1980.

[A16] Unbehauen, H., Editor, "Methods and Applications in Adaptive Control", Proc. of an Int. Symp., Bochum, 1980, in the series "Lecture Notes in Control & Information Sciences", no 24, published by Springer-Verlag (1980).

[A17] Landau, Y.D., "Model Reference Adaptive Control and Stochastic STR - Towards Cross-Fertilization", presented at AFOSR Workshop on Adaptive Control, University of Illinois, Urbana-Champaign, May 1979.

[A18] Egardt, B.O., "Stability of Adaptive Controllers", Lecture Notes in Control and Information Sciences, Vol 20, published by Springer-Verlag (1979).

[A19] Johnson, C.R. Jr., "Input Matching, Error Augmentation, Self-Tuning, and Output Error Identification: Algorithmic Similarities in Discrete Adaptive Model Following", to appear in IEEE Trans. Auto Contr.

[A20] Feuer & Morse, "Local Stability of Parameter-Adaptive Control Systems", Johns Hopkin Conference on Information Science and Systems.

[A21] Peterka & Astrom, "Control of Multivariable Systems with Unknown but Constant Parameters", Proc. 3rd IFAC Symposium on Ident. and System Parameter Estimation, The Hague, June 1973.

[A22] Borisson, U., "S-TR for a Class of Multivariable Systems", Automatica, Vol 15, pp 209-215 (1979).

[A23] Egardt, B.O., "Stability of MRA and STR", Thesis CODEN: LUTFD2/(TFRT-1017)/1-163/Dept. of Auto. Control, Lund Institute of Technology, Lund, Sweden, 1978.

[A24] Astrom, "Self Tuning control of a Fixed Bed Chemical Reactor System", Report CODEN: LUTFD2/(TFRT-3151)/1-066/, Dept. of Auto Control, Lund Institute of Technology, Lund, Sweden, 1978.

[A25] Astrom & Wittenmark, "On the Control of Constant But Unknown Systems", Proc 5th IFAC World Congress, Paris, 1972.

[A26] Parks, P.C., "Lyapunov Redesign of MRAC Systems", IEEE Trans Autom. Control AC-11, pp 362-367 (1966).

[A27] Lindorff & Caroll, "Survey of Adaptive Control Using Lyapunov Design", Int. J. Control 18, pp 897-914 (1973).

[A28] Narendra & Kudva, "Stable Adaptive Schemes for System Identification and Control, Parts I and II", IEEE Trans. Syst. Man Cybern, SMC-4, pp 542-560 (1974).

APPENDIX 2

APPENDIX 2

THE AUTOCORRELATION AND CROSSCORRELATION FUNCTIONS

In general a stationary time series can be described by its mean, variance and autocorrelation function. The calculation of sample estimates of the autocorrelation function is a non-structural approach which provides a first step in the analysis of time series. A priori knowledge of appropriate model order and type of an unknown time series is usually not available. Initial use of the autcorrelation function estimates is therefore necessary to identify the type of model needed.

In the same way that the autocorrelation function is used to identify stochastic models of time series, the data analysis tool applied for the identification of transfer function models is the crosscorrelation function between the input and output time series. In this appendix, definitions of the autocorrelation and crosscorrelation functions will be presented followed by their appropriate estimation using finite time series. Sample calculations will also be provided.

A2.1 The Autocorrelation Function

The autocorrelation function is an useful device for describing the behaviour of stationary processes. A discrete process is said to be strictly stationary if the joint distribution of any set of observations is unaffected by a positive or negative time shift of all the observations.

The joint distribution can be inferred by analyzing pairs of values ((z_t, z_{t+k})) of the time series separated by a constant interval or lag k . The covariance between z_t and its value z_{t+k} separated by k intervals of time is called the autocovariance at lag k and is defined by

$$\gamma_k = \mathrm{cov}[z_t, z_{t+k}] = E[(z_t - \mu)(z_{t+k} - \mu)] \qquad \text{(A2.1)}$$

Similarily, the autocorrelation at lag k is

$$\rho_k = E[(z_t - \mu)(z_{t+k} - \mu)] / (E[(z_t - \mu)^2] E[(z_{t+k} - \mu)^2])^{1/2} \qquad \text{(A2.2)}$$

For a stationary process the variance σ_z^2 is the same at any time t , therefore

$$\rho_k = \gamma_k / \sigma_z^2 = \gamma_k / \gamma_0$$

which implies that $\rho_0 = 1$.

In practice, we have a finite set of observations N, of the time series. Therefore an estimate of the autocovariance and autocorrelation functions is required. Box and Jenkins [1] use the following estimate, generally accepted to be the most satisfactory approximation, of the k lag autocorrelation ρ_k .

$$r_k = c_k / c_0 \qquad (A2.3)$$

where

$$c_k = (1/N) \sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z}) \quad k = 0, 1, 2, \ldots K \qquad (A2.4)$$

is the estimate of the autocovariance γ_k and $\bar{z}$ is the mean of the time series. The following sample calculation is taken from Box and Jenkins [1] page 32.

We now illustrate (A2.4) by calculating r_1 for the first 10 values of the batch data given in Table 2.1. The mean $\bar{z}$ of the first ten values in the table is 51 and so the deviations about the mean are −4, 13, −28, 20, −13, 13, 4, −10, 8, −3.

TABLE 2.1

1–15	16–30	31–45	46–60	61–70
47	44	50	62	68
64	80	71	44	38
23	55	56	64	50
71	37	74	43	60
38	74	50	52	39
64	51	58	38	59
55	57	45	59	40
41	50	54	55	57
59	60	36	41	54
48	45	54	53	23
71	57	48	49	
35	50	55	34	
57	45	45	35	
40	25	57	54	
58	59	50	45	

$$\sum_{t=1}^{9} (z_t - \bar{z})(z_{t+1} - \bar{z}) = (-4)(13) + (13)(-28) + \cdots + (8)(-3)$$

$$= -1497$$

Hence $$c_1 = \frac{-1497}{10} = -149.7$$

Similarly we find that $c_0 = 189.6$. Hence

$$r_1 = \frac{c_1}{c_0} = \frac{-149.7}{189.6} = -0.79$$

A2.2 The Crosscorrelation Function

Transfer function models can be used to represent dynamic systems relating, the simplest case being a single output time series Y_t to a single input series X_t . It is assumed that pairs of observations ((X_t,Y_t)) are available at equispaced intervals of time. This pair of time series is called a bivariate stochastic process.

In the previous discussion, we have seen that a stationary stochastic process can be described by its mean, variance and autocorrelation function or its respective estimates. In general, a bivariate stochastic process need not be stationary. However, the appropriate differenced process is assumed to be stationary. The assumption of stationarity implies constant means and and constant variances σ_x^2 and σ_y^2 . Figure A2.1 shows the different kinds of covariances that need to be considered.

The autocovariance coefficients of each constituent series at lag k are as previously defined. However we now use the following notation $\gamma_{xx}(k)$ and $\gamma_{yy}(k)$ to specify the autocovariances of the X_t and Y_t time series.

$$\gamma_{xx} = E[(x_t - \mu_x)(x_{t+k} - \mu_x)] = E[(x_t - \mu_x)(x_{t-k} - \mu_x)] \qquad \text{(A2.5)}$$

$$\gamma_{yy} = E[(y_t - \mu_y)(y_{t+k} - \mu_y)] = E[(y_t - \mu_y)(x_{y-k} - \mu_y)] \qquad \text{(A2.6)}$$

The other covariances to be considered are the crosscovariance coefficients between X_t and Y_t at lag k and the crosscovariance coefficients between Y_t and

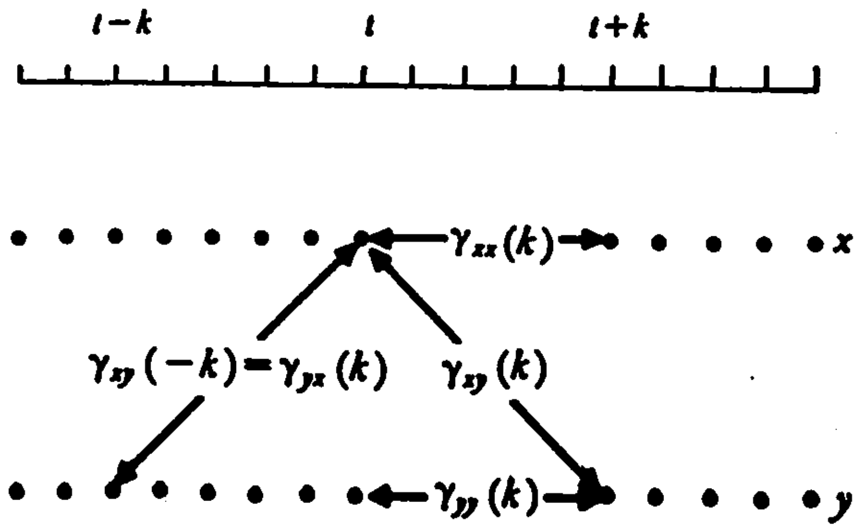


Figure A2.1: Autocovariances and Crosscovariances of a Bivariate Process

X_t at lag k .

$$\gamma_{xy} = E[(x_t - \mu_x)(y_{t+k} - \mu_y)] \qquad k = 0, 1, 2, 3, \ldots \qquad \text{(A2.7)}$$

$$\gamma_{yx} = E[(y_t - \mu_y)(x_{t+k} - \mu_x)] \qquad k = 0, 1, 2, 3, \ldots \qquad \text{(A2.8)}$$

In general $\gamma_{xy}(k)$ and $\gamma_{yx}(k)$ are not equivalent. However since

$$\gamma_{xy}(k) = \gamma_{yx}(-k)$$

we need only define one function $\gamma_{xy}(k)$ for $k = 0, \pm 1, \ldots$. The function $\gamma_{xy}(k)$ is called the crosscovariance function of the bivariate process. The crosscorrelation function of the bivariate process is defined by

$$\rho_{xy}(k) = \gamma_{xy}(k) / \sigma_x \sigma_y \qquad k = 0, \pm 1, \ldots \qquad \text{(A2.9)}$$

Since $\rho_{xy}(k)$ is not in general equal to $\rho_{xy}(-k)$, the crosscorrelation function, unlike the autocorrelation function, is not symmetric around $k = 0$.

An estimate of the crosscovariance coefficient at lag k using the suitably differenced time series is provided by

$$c_{xy}(k) = \begin{cases} (1/N) \sum_{t=1}^{N-k} (x_t - x)(y_{t+k} - y) & k = 0, 1, 2, \ldots K \\ (1/N) \sum_{t=1}^{N-k} (y_t - y)(x_{t-k} - x) & k = 0, -1, \ldots -K \end{cases} \qquad \text{(A2.10)}$$

where x , y are the respective means of x and y . The estimate $r_{xy}(k)$ of the crosscorrelation coefficient $\rho_{xy}(k)$ at lag k may be obtained by substitution of the estimates $C_{xy}(k)$ for $\gamma_{xy}(k)$, $s_x = (C_{xx}(0))^{1/2}$ for σ_x^2 and $s_y = (C_{yy}(0))^{1/2}$ for σ_y^2

yielding

$$r_{xy}(k) = C_{xy}(k) \,/\, S_x S_y \qquad k = 0, \pm 1, \ldots \qquad \text{(A2.11)}$$

The following example is taken from Box and Jenkins [1] page 374.

An example. In practice we would need at least 50 pairs of observations to obtain a useful estimate of the cross correlation function. However, to illustrate the formulae (A2.10) and (A2.11) we compute an estimate of the cross correlation function at lags +1 and −1 for the following series of 5 pairs of observations

t	1	2	3	4	5
x_t	11	7	8	12	14
y_t	7	10	6	7	10

Now $\bar{x} = 10.4$, $\bar{y} = 8$, so that the deviations from the means are

t	1	2	3	4	5
$x_t - \bar{x}$	0.6	−3.4	−2.4	1.6	3.6
$y_t - \bar{y}$	−1.0	2.0	−2.0	−1.0	2.0

Hence

$$\sum_{t=1}^{4} (x_t - \bar{x})(y_{t+1} - \bar{y}) = (0.6)(2.0) + (-3.4)(-2.0) + (-2.4)(-1.0) + (1.6)(2.0) = 13.60$$

and

$$c_{xy}(1) = 13.60/5 = 2.720$$

Using $s_x = 2.577$, $s_y = 1.673$, we obtain

$$r_{xy}(1) = \frac{c_{xy}(1)}{s_x s_y} = \frac{2.720}{(2.577)(1.673)} = 0.63$$

Similarly $\sum_{t=1}^{4} (y_t - \bar{y})(x_{t+1} - \bar{x}) = -8.20$. Hence, $c_{xy}(-1) = -1.640$ and

$$r_{xy}(-1) = \frac{-1.640}{(2.577)(1.673)} = -0.38$$

APPENDIX 3

APPENDIX 3

THE PREWHITENING OPERATOR

The preliminary identification of transfer function-noise models is performed using the estimated crosscorrelation function of the suitably differenced input and output series. Suppose the model relating Y_t and X_t may be written in the following form

$$y_t = v_0 x_t + v_1 x_{t-1} + v_2 x_{t-2} + \dots + n_t \qquad (A3.1)$$

where $y_t = \nabla^d Y_t \quad x_t = \nabla^d X_t \quad n_t = \nabla^d N_t$

are stationary processes with zero means. Multiplication of equation (A3.1) by x_{t+k} for $k \geq 0$ yields

$$x_{t+k} y_t = v_0 x_{t+k} x_t + v_1 x_{t+k} x_{t-1} + \dots + x_{t+k} n_t \qquad (A3.2)$$

If the further assumption that x_{t+k} is uncorrelated with n_t is made then taking expectations of equation (A3.2) yields the set of equations

$$\gamma_{xy}(k) = v_0 \gamma_{xx}(k) + v_1 \gamma_{xx}(k-1) + \dots \quad k = 0, 1, 2, 3, \dots \qquad (A3.3)$$

Assuming that the weights v_j are essentially zero beyond $k = K$, then the first $K+1$ of the equations (A3.3) can be written as

$$\gamma_{xy} = \Gamma_{xx} \, v \qquad (A3.4)$$

where

$$V = \begin{bmatrix} v_0 \\ v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_K \end{bmatrix} \qquad \gamma_{xy} = \begin{bmatrix} \gamma_{xy}(0) \\ \gamma_{xy}(1) \\ \cdot \\ \cdot \\ \cdot \\ \gamma_{xy}(K) \end{bmatrix} \qquad \Gamma_{xx} = \begin{bmatrix} \gamma_{xx}(0) & \gamma_{xx}(1) & \dots & \gamma_{xx}(K) \\ \gamma_{xx}(1) & \dots & & \cdot \\ \cdot & \dots & & \cdot \\ \cdot & \dots & & \cdot \\ \gamma_{xx}(K) & \dots & & \gamma_{xx}(0) \end{bmatrix}$$

Using the respective estimates of the autocorrelation and crosscorrelation functions in equation (A3.4) provides $K+1$ linear equations for the first $K+1$ weights. However these equations are cumbersome to solve and do not in general provide efficient estimates. Furthermore the knowledge of the value of K is required.

Considerable simplication in the preliminary identification would be achieved if the input series to the system were white noise. However, a white noise input series is not always available. When the original input series follows some other stochastic process, simplication is possible by prewhitening.

Assuming that the suitably differenced input process is stationary and capable of representation by some member of the general linear class of autoregressive moving average models. Then given the data, a model for the x_t process may be obtained

$$\phi_x(B)\,\theta_x^{-1}(B)\,x_t = \alpha_t \qquad \text{(A3.5)}$$

which, to a close approximation, transforms the correlated input series x_t into an uncorrelated white noise series α_t .

If the same transformation is now applied to y_t we obtain

$$\beta_t = \phi_x(B)\theta_x^{-1}(B)y_t \quad \text{(A3.6)}$$

Then the model (A3.1) may be written as

$$\beta_t = v(B)\alpha_t + \epsilon_t \quad \text{(A3.7)}$$

where ϵ_t is defined as the transformed noise series

$$\epsilon_t = \phi_x(B)\theta_x^{-1}(B)n_t \quad \text{(A3.8)}$$

Multiplying both sides by α_{t-k} and then taking expectations yields

$$\gamma_{\alpha\beta}(k) = v_k\sigma_\alpha^2 \quad \text{(A3.9)}$$

Therefore, after prewhitening of the input, the crosscorrelation function between the prewhitened input and similarily transformed output is directly proportional to the impulse response function. The effect of prewhitening is to convert the non-orthogonal set of equations (A3.4) into the orthogonal set (A3.9).

The preliminary estimates v_k obtained using the prewhitened operator provide a rough basis for selecting suitable operators $\delta(B)$ and $\omega(B)$ in the transfer function model.

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA