IDENTIFICATION OF SYSTEMS FROM MULTIRATE DATA

MAY SU TUN

NATIONAL UNIVERSITY OF SINGAPORE

2004

IDENTIFICATION OF SYSTEMS FROM MULTIRATE DATA

MAY SU TUN

(B.Sc (Honours) I.C YU, Yangon), (B.E (Chemical) YTU, Yangon)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ENGINEERING

DEPARTMENT OF CHEMICAL & BIOMOLECULAR ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2004

ACKNOWLEDGEMENTS

First and foremost, I would like to express and record my deepest gratitude and indebtedness to my supervisor Dr. Lakshminarayanan Samavedham for his sincere and kind support, guidance and encouragement throughout this research work. I greatly thank him for system identification lectures and advices. Other than technical things, I can learn good aspects of human mind from his generous and warm nature. Due to his kind and positive attitude, I am able to make a contribution to this extent. I would like to express appreciation to him again in preparation of this manuscript and above all his understanding and help in different ways, all the time. I would like to express my deep appreciation and gratefulness to my co-supervisor Dr. Arthur Tay for his advices, patience, understanding, and providing me the freedom to perform this work.

I really appreciate my colleagues Kyaw Tun, Madhukar, Prabhat, Mranal, Dharmesh, Rampa, Balaji, and Rohit for their help and friendship which resulted in a fine working environment. My close friends Thet Su Hlaing, Ne Lin, Mya Mya Khin, Khin Yin Win, Khin Moh Moh Aung and other friends who are not explicitly expressed by name here are also much appreciated for their inspiration and healthy friendship. I dedicate this work to my beloved ones; my parents, brother and sister - special thanks to them for their moral support. Lastly, my sincere thanks to the National University of Singapore for this educational opportunity and for providing support in the form of a research scholarship.

I asked for WISDOM,
And GOD gave me problems to solve
…….. Anonymous

TABLE OF CONTENTS

SUMMARY

Multirate systems are very common in the chemical industries where the measurements of variables such as compositions, melt flow index, molecular weight distribution are available infrequently while that of variables such as temperature, flow rate, pressure are measured frequently. Utilizing infrequent measurements of the controlled variables *alone* in the control strategy will naturally lead to poor quality products or suboptimal process operation. It would naturally be advantageous to develop "fast rate" process models by bringing together the "fast" (frequent) and "slow" (infrequent) measurements and use it for applications such as process control and soft sensing. The availability of fast-rate model is advantageous for any model based control strategy including Model Predictive Control (MPC). Many identification methods are developed and applicable for the identification of single-rate system in which the sampling interval of input variables and output variables are identical. The topic of multirate system identification was developed very little in the past. The missing data during the infrequent sampling interval were estimated conventionally using linear interpolation, cubic interpolation, zero order hold etc. With such naïve approximations, the estimated models tend to be of poor quality and result in deteriorated controller performance.

To alleviate this problem, a technique known as "lifting" has been applied in the recent past to enable the identification of fast rate process models from multirate data. In this technique, the fast sampled input data are "lifted" (using a lifted operator) to generate a slow-rate multi-input sequence (each fast sampled input variable is lifted into several slow rate input sequences). For the non-integer ratio of sampling interval, both input and output channels are lifted with proper lifting operator into a slow-rate

system with common period. Then, any multivariable system identification method such as the popular **s**ub**s**pace based **s**tate **s**pace **id**entification methods (4SID methods) are employed for the identification of the lifted slow-rate model. The fast-rate model is subsequently extracted from the identified slow-rate system using one of the several available approaches. The lifting technique considered here can handle the regularly sampled data system only (i.e. multirate but regularly sampled data).

In a regression based method named **d**ata **s**election **a**nd **r**egression (DSAR) method, the fast sampled inputs and the slow sampled process outputs are stacked into appropriate matrices. The model is then determined using ordinary least squares. For highly correlated data, methods such as principal component regression (PCR) or partial least squares (PLS) may be applied. The obtained model is similar to the finite impulse response (FIR) model and is non-parsimonious. This model may then be compacted if there is a need. The DSAR method is applicable to irregularly sampled data and also in situations where data is sampled very infrequently. The evaluation of this method on industrial data is also reported in this thesis.

The effect of different kinds of input signals on these methods (lifting and DSAR) is also studied. The ratio of sampling intervals (denoted by γ) could vary from 1 to a large number and this could affect the quality of the identified model. Thus, the effect of γ to the identified model was also studied. Besides these, nonlinear multirate system identification methods are developed. Some of the chemical processes such as heat exchangers, distillation units and pH neutralization process which have nonlinear behavior can be represented by the Hammerstein or Wiener model. Thus, the nonlinear identification methods for Hammerstein model and Wiener model from multirate sampled data are developed. The application of the developed method is evaluated with both simulated and experimental data.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Overview of System Identification

Often, systems or subsystems cannot be modeled based on physical insights; because the function of the system or its construction is unknown or it would be too complicated to sort out the physical relationship. In such situations, the mathematical model of the process can only be obtained empirically. This is the topic of system identification. System identification is the mathematical modeling of a dynamic system from test or experimentally measured input/output data set. The dynamic system is one in which the current output value depends not only on the current external stimuli but also on their earlier values. Zadeh (1962) defined system identification as: *the determination on the basis of input and output, of a system (model) within a specified class systems (models), to which the system under test is equivalent (in terms of a criterion).* System identification is widely used in many fields such as process industries, economics, biomedical and many other fields of science.

Advanced control technology or model-based control system design relies heavily on reasonably accurate process models. This has been the case since the birth of 'modern control theory' in the early 1960s. Based on the models obtained from system identification, advanced model based control technologies such as Model Predictive Control (MPC) have been successfully applied in the chemical process industries. Moreover, identified models are widely used for fault detection, pattern recognition, adaptive filtering, linear prediction and other purposes.

In process industries, the process outputs are driven by the input variables (manipulated variables and disturbances). The measured input and output variables provide useful information about the system. Process/Control engineers try to model chemical processes by collecting the input/output data after subjecting the process to open loop or closed loop identification tests. In the open loop test, there is no feedback controller and the test signals are the process input signals; in the closed loop test, the test signal is added at the set point. Compared to the open loop, closed loop identification is more difficult because the input is correlated with the disturbance due to feedback. This thesis concentrates exclusively on open loop identification. The effect of input signal on the different identification methods is explored.

In the early days of the control technology, analog control based on continuous models was employed. Later, and almost exclusively these days, discrete domain models are widely used. This is due to the deployment of computer process control systems which are based on measurements made at discrete time instants (i.e. sampled data control systems). System identification techniques for linear systems are well established and have been widely applied. Most often, an MPC controller uses a linear dynamic model of the process that is obtained by the way of black-box identification. However, most of the chemical processes are nonlinear (e.g. heat exchanger, pH neutralization process, distillation column, waste water treatment plant, bioreactor). Most processes encountered in practice are nonlinear to some extent. Although it may be possible to represent systems which are perturbed over a restricted operating range by a linear model, in general, nonlinear process can only be adequately characterized by a nonlinear model. Because of these reasons, this thesis focuses on discrete models only but covers both linear and nonlinear models.

System identification is done by adjusting the parameters of a chosen model until its output coincides as much as possible with the measured output. For parametric models, it is necessary to specify the structure. Well known model parameterizations include models such as AutoRegressive (AR) model, AutoRegressive eXogeneous (ARX) model, AutoRegressive Moving Average (ARMA) model, AutoRegressive Moving Average eXogeneous (ARMAX) model, Box-Jenkins (BJ) model and Output Error (OE) models. In addition, state space models are also well established and are extensively used due to their convenience in representing multivariable process. For linear systems, nonparametric models include the finite impulse response (FIR) models, step response models (these models can be obtained using correlation analysis) and the frequency domain representation (Bode/Nyquist plot).

Model identification is essentially an iterative procedure that involves choosing a model structure, plant experimentation (that is commensurate with the chosen model structure and one that meets the operational constraints), parameter estimation and model validation. The iterative procedure may also involve choosing a different and complex model structure should the simpler models prove to be ineffective in explaining the observed experimental data. If the linear model structures mentioned above is not sufficient in describing the system, the suitability of nonlinear model structures need to be investigated. There are several ways to describe the nonlinearity of systems. The Volterra series was originally developed to describe the nonlinearity of a very general class of nonlinear time-invariant process. Although the Volterra series representation of nonlinearity provides theoretical understanding of nonlinearity, the number of coefficients in this model is excessive and places enormous requirements on the identification procedure (quality and quantity of data). Alternate representations for nonlinear processes include the Wiener model (a model

in which a linear dynamic block is followed by a nonlinear static block) and the Hammerstein model in which nonlinear zero-memory gain is followed by a linear dynamic part (reverse of the Wiener model). These two models are among the well known block-oriented models - bases on these models, many other block-oriented models like Hammerstein-Wiener (N-L-N) model, L-N-L model and more complex parallel connection of these described models are developed. The identification of a block-oriented nonlinear model is more difficult than that of a linear model because nonlinear model identification needs a richer probing (input) signal and a robust identification procedure (as it may involve iterative solution or nonlinear optimization).

Billings and Voon (1986) described a popular discrete-time model, Nonlinear ARMAX (NARMAX) model, in which they introduced a nonlinear function term to the ARMAX model. Other model structures are Nonlinear Moving Average models with eXogeneous inputs (NMAX), Nonlinear AutoRegressive models with eXogeneous inputs (NARX) and the Nonlinear Additive ARX (NAARX) model. Like in the linear case, the selection of appropriate model structure is important in nonlinear identification. Hammerstein and Wiener models are widely used because of their adequacy of representing the many chemical processes that are nonlinear in nature. Because of their usefulness in identification of nonlinear chemical system, this thesis tries to explore the identification of these two models.

1.2 Multirate System and Multirate Identification

Different from single rate systems in which the inputs and outputs are measured at the same sampling interval, multirate systems are sampled-data systems with non-uniform

sampling intervals. Multirate systems are very common in chemical process industries in which different variables are sampled at different rates. In process units such as distillation columns and reactors, variables such as temperature, pressure, flow rate, etc. can be measured frequently while variables such as composition, molecular weight distribution, melt flow index etc. are obtained infrequently. This is because measurements of the latter type variables often involve elaborate offline analysis. These measurements are obtained once in several minutes or even once in several hours. These features naturally lead to a multirate system.

Theoretically, there are different ways of process modeling - first principles model (arising out of mass, energy and momentum balances), black box models (empirically developed using observed process data) or gray-box model (where the first principles model contains terms that are fitted using a black box approach). This thesis examines the black box modeling approach only. This is because of the fact that the input-output measurements are readily available from plant historical databases or from carefully designed process experiments. Black box models lend themselves more easily for applications such as controller design or output predictions. Most of the successful system identification methods in both transfer function domain and state space domain can only be applied to single-rate input/output data. Very few algorithms have been developed for identification of process models from Multirate input/output data. Conventionally, engineers interpolate the inter-sample input/output from the slowly sampled measurements and then estimate fast-rate model based on the interpolated data set. The model obtained from such *ad hoc* interpolation techniques cannot capture the actual process dynamics very well (and particularly when the ratio of sampling intervals becomes large). This situation provides the motivation to investigate multirate system identification procedures.

Verhaegen and Yu (1995) presented a technique to estimate the lifted model (the concept of lifting will be explained in Chapter 3) of Multirate system in the state-space domain. They represented Multirate system as a periodic system and estimated the lifted model with the multivariable output error state space method. Their method cannot handle the crucial constraint, causality constraint, in identification of lifted models. Li et al. (2001) made some modification on their work to overcome the causality constraint – with this modification, most of the existing identification algorithms can be applied for identification of lifted system (slow model). After that, Li tried to extract fast rate model using two approaches. Wang et al. (2004) improved upon Li's work in the extraction of the fast rate model. Identification of the slow rate model is accomplished using state space methods that are able to effectively handle multivariable processes. It is important to note that all of these works deal with linear systems only.

Gopaluni et al. (2003) explored a Multirate identification algorithm in which they used an iterative procedure. They first identified an FIR model from the Multirate data. Based on this model, the missing data points in the slow sampled measurement are estimated using the expectation maximization approach. Then they identified a new model iteratively using the estimated missing data points and original data set until the models converge. Their method is applicable to irregularly sampled data system as well. Lakshminarayanan (2000) developed Data Selection and Regression (DSAR) method for the identification of multirate system. The advantages of his work is not only it is able to handle the large ratio of sampling interval it is also useful to irregularly sampled data system. This method is applicable to chemical industry in which the ratio of sampling intervals is very large.

1.3 Scope and Organization of the Thesis

This thesis deals with discrete data only and focuses on Multirate system identification using the lifting and DSAR methods. We consider both linear and nonlinear systems. The effect of different kinds of input signal and the effect of the ratio of sampling intervals are studied using simulated case studies. We explore nonlinear multirate system identification methods for Hammerstein and Wiener models. The evaluations of these techniques are provided with simulated case studies. The best excitation signal for the identification of these models is proposed. The industrial application of DSAR method and development of a soft sensor are evaluated with industrial data set. The organization of the thesis is as follows. Chapter 2 introduces subspace models identification using 4SID methods. The subspace identification methods are used extensively in the rest of the thesis. The working examples of subspace based state space identification methods are demonstrated through case studies involving single rate data. Two multirate identification methods are described in Chapter 3 and 4 of the thesis respectively. Chapter 3 introduces the readers to a method called “Lifting”. Using the lifting technique, we demonstrate the identification of a slow rate model which is then converted to a fast rate model. In Chapter 4, we discuss a method called data selection and regression (DSAR) for the identification of process models from multirate data. Both of the identification approaches are illustrated using suitable examples. In Chapter 5, we provide extensive case studies for Multirate identification - besides simulation examples, we demonstrate Multirate identification using data from laboratory systems as well as from an industrial reactor. Chapter 6 summarizes the contributions of this thesis and makes recommendations for future work.

CHAPTER 2

SUBSPACE-BASED IDENTIFICATION METHODS

2.1 Introduction

Subspace-based identification methods are most suited to identify models in state space form for representing multivariable systems. So, subspace-based identification methods are very useful in the identification of chemical processes. These methods firstly estimate the states directly from the input/output data using linear algebra (QR decomposition or singular value decomposition or generalization of these methods) and then figure out the state space model matrices (A, B, C, D) using the least squares method. It is possible to obtain more efficient model with a smaller number of regressors by using a state space structure. The states produced by these approaches are not real states; these states are not physically meaningful. They are optimal linear combination of past inputs and outputs of the plant. In subspace identification algorithms, the only one parameter needed to specify is the order of the system. The optimal model order can be determined by Akaike Information Criterion (AIC) or by inspection of certain singular values. Subspace identification algorithms not only guarantee the convergence but also the numerical stability because they are non-iterative and involve the well known linear algebra. A number of subspace identification methods have been developed over the last fifteen to twenty years. A powerful method called Canonical Variate Analysis (CVA) was developed by Larimore in 1990. Starting from 1992, Verhaegen developed Multivariable Output-Error State sPace (MOESP) methods in a series of papers (Verhaegen and Dewilde (1992a,b), Verhaegen (1993, 1994), Verhaegen and Xu (1995), Verhaegen and

Westwick (1996)). Van Overschee and De Moor (1994) developed yet another variant of 4SID methods namely the N4SID method which has been incorporated into the System Identification Toolbox of MATLAB. In this chapter, we present the above mentioned three subspace identification algorithms briefly and then illustrate the application of these methods in the identification of single rate systems using data from experiments and simulations.

2.2 CVA

Larimore's Canonical Variate Analysis (CVA) is a powerful identification tool for linear systems. It can identify correct or close to correct model order even for small sample sizes, low signal to noise ratio or for any choice of probing signals. CVA is based on the Generalized Singular Value Decomposition (GSVD) theory. The optimal memory length and state order are determined using AIC. The estimation of states from input and output data is performed using one of the multivariate techniques, Canonical Correlation Analysis (CCA). CVA estimates are as asymptotically efficient as the maximum likelihood (ML) estimates. In this thesis, the CVA algorithm developed by Lakshminarayanan (1997) is used substantially in the identification of processes. This algorithm is described briefly in this section. Before presenting the CVA algorithm, an important component of it, namely the CCA technique, is introduced.

2.2.1 Canonical Correlation Analysis

Let us have two sets of variables; a set of several predictor variables X and a set of one or more dependent variables Y.

where the size of matrix X is $(ns\ by\ nx)$,

the size of matrix Y is $(ns\ by\ ny)$,

and the rank of X : $rx = \min(ns, nx)$.

Then we define the canonical variates t_1 and u_1.

Canonical Variate in X space is

$$t_1 = X\ j_1 \tag{2.1}$$

Canonical Variate in Y space is

$$u_1 = Y\, l_1 \tag{2.2}$$

Correlation between t_1 and u_1 = $p(t_1, u_1) = \dfrac{j_1^T X^T Y l_1}{\sqrt{j_1^T X^T X j_1}\sqrt{l_1^T Y^T Y l_1}}$ (2.3)

$$= \frac{j_1^T \sum_{XY} l_1}{\sqrt{j_1^T \sum_{XX} j_1}\sqrt{j_1^T \sum_{YY} l_1}} \tag{2.4}$$

Here, p is referred to as the canonical correlation.

The objective is to maximize $\dfrac{j_1^T \sum_{XY} l_1}{\sqrt{j_1^T \sum_{XX} j_1}\sqrt{j_1^T \sum_{YY} l_1}}$. (2.5)

subject to the constraints

$$j_1^T \sum_{XX} j_1 = 1 \tag{2.6}$$

and $$l_1^T \sum_{YY} l_1 = 1 \tag{2.7}$$

The solution can be obtained as

$$j_1 = \sum_{XX}^{-1/2} * (\text{first left singular vector of } \sum_{XX}^{-1/2} \sum_{XY} \sum_{YY}^{-1/2}) \tag{2.8}$$

$$l_1 = \sum_{YY}^{-1/2} * (\text{first left singular vector of } \sum_{YY}^{-1/2} \sum_{YX} \sum_{XX}^{-1/2}) \tag{2.9}$$

Other components can be defined as

$$j_2 = \sum_{XX}^{-1/2} * (\text{second left singular vector of } \sum_{XX}^{-1/2} \sum_{XY} \sum_{YY}^{-1/2}) \tag{2.10}$$

$l_2 = \sum_{YY}^{-1/2}$ * (second left singular vector of $\sum_{YY}^{-1/2} \sum_{YX} \sum_{XX}^{-1/2}$). (2.11)

t_1 is the best predictor in the X space and u_1 is the most easily predicted linear combination in the Y space. The next best linear combination pair (Xj_2, Yl_2), orthogonal to the t_1, u_1 pair, is obtained by using j_2 and l_2. Similar arguments hold for other pairs as well. In this algorithm, each canonical variate is orthogonal to all the previously generated ones. A maximum of $\min(rx, ny)$ canonical variate pairs can be generated.

2.2.2 Canonical Variate Analysis

Consider a system with p inputs and q outputs. We assume that N input/output samples are available.

Consider the following state space model structure in discrete domain

$$X_{t+1} = \Phi X_t + GU_t + W_t$$

$$Y_t = HX + AU_t + BW_t + V_t \quad (2.12)$$

where W_t is state noise and $BW_t + V_t$ is measurement noise. The presence of BW_t in the output equation allows for correlation between state noise (W_t) and the measurement noise $(BW_t + V_t)$. This makes CVA to be compatible to the experimental data that are rich in noise. Our objective is to estimate the (Φ, G, H, A, B, Q, R) matrices (state space matrices). Φ, G, H, A and B are called as system matrices; Q and R are the covariance matrices for W_t and V_t respectively.

Generally, we can define the basic steps in CVA as follows:

- specification of data and maximum memory length
- determine the optimal memory length

- computation of the states using CCA
- choosing the optimal number of states using AIC
- generating the system matrices and estimates for the noise covariance matrices.

Firstly, we can specify the optimal memory length, L using *a priori* knowledge or have to specify the maximum memory length (L^*). L^* must be equal to or greater than the maximum possible delay plus 2.

We can then determine the optimal memory length using some methods e.g. Auto Regressive (AR) modeling or by applying augmented upper diagonal identification (AUDI) in which the optimal model order is the optimal memory length.

We can define the past space, P and future space, F as follows:

At each time instant k,

$$P_k = [Y_{k-1}, Y_{k-2}, \cdots, Y_{k-L}, U_{k-1}, U_{k-2}, \cdots, U_{k-L}] \tag{2.13}$$

$$F_k = [Y_k, Y_{k+1}, \cdots, Y_{k+L-1}] \tag{2.14}$$

where

$$Y = [Y_1, Y_2, ..., Y_q],$$

$$U = [U_1, U_2, ..., U_p],$$

and $k = [L+1, L+2, \cdots, N-L+1]^T$.

Then, by stacking up the P_k's and F_k's, we can construct the past and future spaces (P and F matrices) respectively.

Thirdly, we relate the past and future spaces using CCA. The canonical variates of the past space are the pseudostates.

$$X_i = Pj_i \tag{2.15}$$

By this way, a total of $\min(pL, qL)$ states can be generated.

Next, the optimal model order (optimal number of states) is chosen using AIC (with small sample correction factor).

$$AIC_k = (N - 2L + 1)\,(q(1 + ln(2\pi)) + ln\left|\sum\nolimits_{ee}^{k}\right| + 2\delta_k M_k \tag{2.16}$$

where

AIC_k = AIC for model order k $\qquad (k = 1,2,\ldots,L)$

δ_k = small sample correction factor

$$= \frac{N}{N - \left(\frac{M_k}{q} + \frac{q+1}{2}\right)} \tag{2.17}$$

M_k = number of independent parameters in the k^{th} order state space model

$$= 4kq + q(q+1) + 2kp + 2qp \tag{2.18}$$

$\sum_{ee}^{k}$ = error covariance matrix for model order k

$$= \frac{1}{N - 2L + 1} \sum_{t=L+1}^{N-L+1} \left(y(t) - \hat{y}^k(t)\right)^T \left(y(t) - \hat{y}^k(t)\right) \tag{2.19}$$

Finally, we generate system matrices and noise covariance matrices as follows:

System matrices can be estimated as:

$$\begin{bmatrix} \Phi & G \\ H & A \end{bmatrix} = \begin{bmatrix} J_k^T \sum_{p(t+1)p(t)} J_k & J_k^T \sum_{p(t+1)u(t)} \\ \sum_{y(t)p(t)} J_k & \sum_{y(t)u(t)} \end{bmatrix} \begin{bmatrix} J_k^T \sum_{p(t)p(t)} J_k & J_k^T \sum_{p(t)u(t)} \\ \sum_{u(t)p(t)} J_k & \sum_{y(t)u(t)} \end{bmatrix}^{-1} \tag{2.20}$$

Noise covariance matrices can be generated as:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} J_k^T \sum_{p(t)p(t+1)} J_k & J_k^T \sum_{p(t+1)y(t)} \\ \sum_{y(t)p(t+1)} J_k & \sum_{y(t)y(t)} \end{bmatrix} - \Psi \tag{2.21}$$

with

$$\Psi = \begin{bmatrix} \Phi & G \\ H & A \end{bmatrix} \begin{bmatrix} J_k^T \sum_{p(t)p(t+1)} J_k & J_k^T \sum_{p(t)y(t)} \\ \sum_{u(t)p(t+1)} J_k & \sum_{u(t)y(t)} \end{bmatrix} \tag{2.22}$$

where

$$J_k = [j_1 \mid j_2 \mid j_3 \mid \cdots \mid j_k]$$

Each j_i is the weight vector corresponding to the canonical variate i.e. $X_i = Pj_i$.

$$B = S_{21}S_{11}^{\dagger} \quad (2.23)$$

$$Q = S_{11} \quad (2.24)$$

$$R = S_{22} - S_{21}S_{11}^{\dagger}S_{12}, \quad (2.25)$$

where † indicates the pseudoinverse operation.

In the above expressions,

$$p(t\} = [y(t-L)\ u(t-L) \ldots y(t-2)\ u(t-2)\ y(t-1)\ u(t-1)] \quad (2.26)$$

$$y(t) = [y_1(t)\ y_2(t) \ldots y_q(t)] \quad (2.27)$$

$$u(t) = [u_1(t)\ u_2(t) \ldots u_p(t)] \quad (2.28)$$

$$p(t+1) = [y(t+1-L)\ u(t+1-L) \ldots y(t-1)\ u(t-1)\ y(t)\ u(t)] \quad (2.29)$$

and $\sum$ signifies the covariance matrices. The predictions of the k^{th} order state space is given by

$$\hat{y}^{\,k}(t) = \sum_{y(t)p(t)} J \sum_k (J_k^T \sum_{p(t)p(t)} J_k)^{-1} J_k^T p^T(t) \quad (2.30)$$

The computation of the prediction error series and its covariance matrix is now straightforward.

2.3 N4SID

Van Overschee and De Moor (1991a, 1991b) developed a class of algorithms for the identification of state space models. Their method is called N4SID. Their algorithm is similar to that of Moonen et al. (1989) for the purely deterministic case. The state

sequence is constructed by projecting the input-output data (containing both deterministic and stochastic parts) in which future output is projected to past and future input and past output. Then the state space matrices are estimated from the constructed state sequence using least squares prediction. N4SID algorithms guarantee convergence because there are no iterative calculations and because no nonlinear optimization is involved. Besides, these N4SID algorithms are numerically stable since they use only QR and singular value decomposition methods. The model order is determined from non-zero singular values (details can be found in Van Overschee and De Moor (1994)).

2.4 MOESP

MOESP stands for **M**ultivariable **O**utput **E**rror **S**tate s**P**ace identification method. MOESP was developed by Verhaegen and Dewilde (1992a). In their algorithm, the constructed input-output Hankel matrices are pretreated by QR factorization and then singular value decomposition (SVD) is performed. The matrices resulting from QR factorization, which has the same column space of extended observability matrix, is treated by SVD and then from the resulting matrices, Φ and H state matrices are estimated. In the second stage, the G and A state matrices are determined. This is different from the earlier methods where all the system matrices are estimated simultaneously (in a single step). The model order is determined by number of nonzero singular values. MOESP is also mathematically stable and guarantees convergence.

In the next section, we will illustrate the identification of processes using single rate data.

2.5 Application of CVA, N4SID, MOESP on single rate data

2.5.1 Experimental Examples

2.5.1.1 Case Study I

The experimental data obtained from the stirred tank heater which is set up in University Of Alberta (Canada) is used for identification study. These data were downloaded from the University of Alberta (Computer Process Control (CPC) group, Department of Chemical and Materials Engineering) website. The process is computer controlled with cold water valve position being the manipulated variable and the water level in the tank as the output. An open-loop experiment was performed. These quantities are measured in units of current and they have linear relationship with their respective physical units. Cold water valve position is perturbed once every 40 seconds using low frequency random binary sequence (RBS) signal and the tank water level is also sampled once every 40 seconds. The three different models estimated by CVA, N4SID, and MOESP approach are cross validated by comparison with the measured output data. It can be observed that CVA, N4SID and MOESP identification methods can adequately identify the single rate linear system from the following cross validation figures (Figure 2.1 to 2.3).

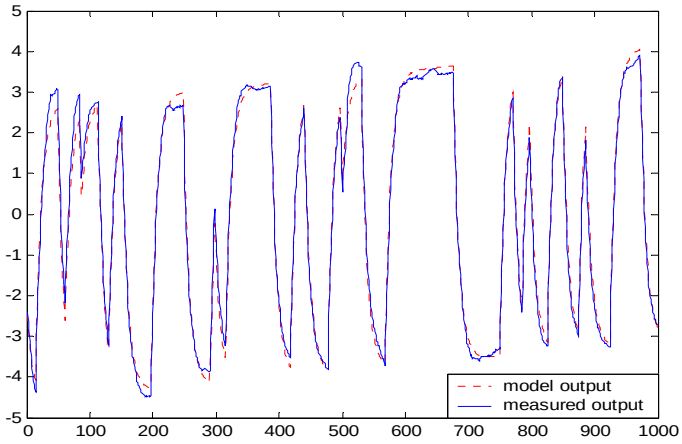

Figure 2.1: Comparison of model output and measured output data using CVA, C1

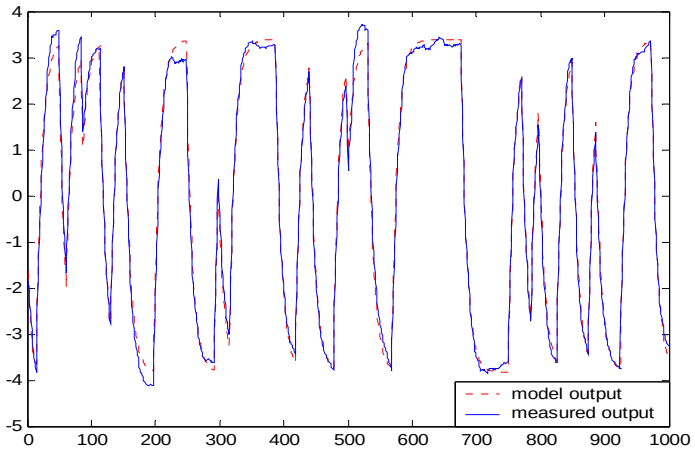

Figure 2.2: Comparison of model output and measured output data using N4SID, C1

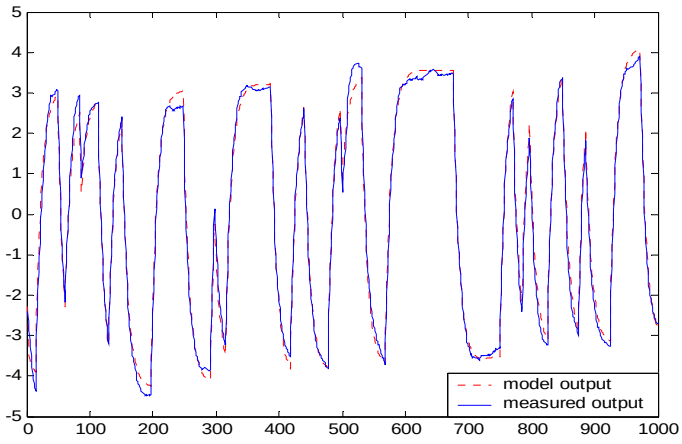

Figure 2.3: Comparison of model output and measured output data using MOESP, C1

2.5.1.2 Case Study II

In this case study, we used the experimental steam-water heat exchanger data obtained from Eskinat et al. (1991). In this example, process water flow rate and water exit temperature are collected as process input and output data respectively. The sampling interval is 12 seconds. Pseudo Random Binary Sequence (PRBS) tests were performed on the heat exchanger. The details of the process nature and operating conditions are available from the above mentioned paper. The process becomes nonlinear when the process is running at constant steam flow rate and at high cool water flow rate because flooding decreases the heat transfer area and heat transfer rate. However, we try to identify the model with linear identification methods namely CVA, N4SID, and MOESP in order to test the appropriateness of the linear identification approach. Figures 2.4, 2.4 and 2.6 show there is nonlinearity (observe the gain mismatch) but the employed three linear subspace methods can adequately identify the mid to high frequency characteristics of the process. The identification of this process data with nonlinear identification method is shown in a later chapter.

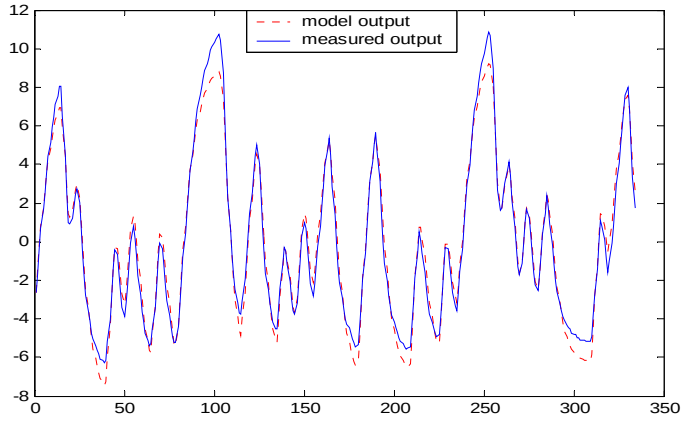

Figure 2.4: Comparison of model output and measured output data using CVA, C2

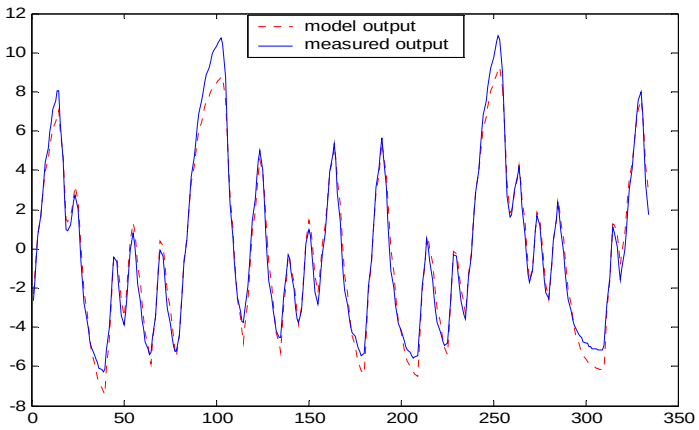

Figure 2.5: Comparison of model output and measured output data using N4SID, C2

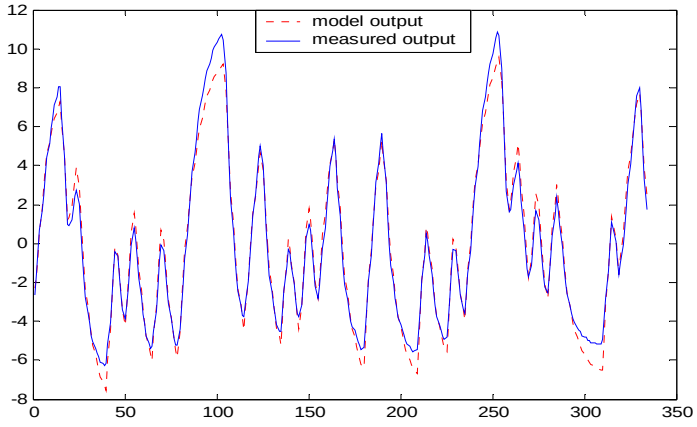

Figure 2.6: Comparison of model output and measured output data using MOESP, C2

2.5.1.3 Case Study III

In this example, we identify an empirical process model for the experimental system available in our research group (Data Analysis and Control System (DACS) group). The schematic of the experimental equipment is shown in Figure (2.7). This experimental set up has three tanks (two tanks of uniform cross section and one tank with a conical base) plus a reservoir. All tanks have heating equipment and the stirrers keep the tank water temperature constant throughout the tank. All tanks are connected with winding pipes for the purpose introducing time delays. In this case study, we concentrated on input and output data of tank 1 - heating power is the input and water temperature is the output. The input was designed as multilevel and multifrequency signal. Input and output are sampled at every one second and the system was identified with the three subspace methods considered here. Validation of these data was performed similar to that of Case Study I. The results of validation (Figure 2.8 through Figure 2.10) show that the system is linear to considerable extent and the three identification methods do perform well over the range of operation. The nonlinearity of this system and nonlinear identification of this data will be discussed in a later chapter.

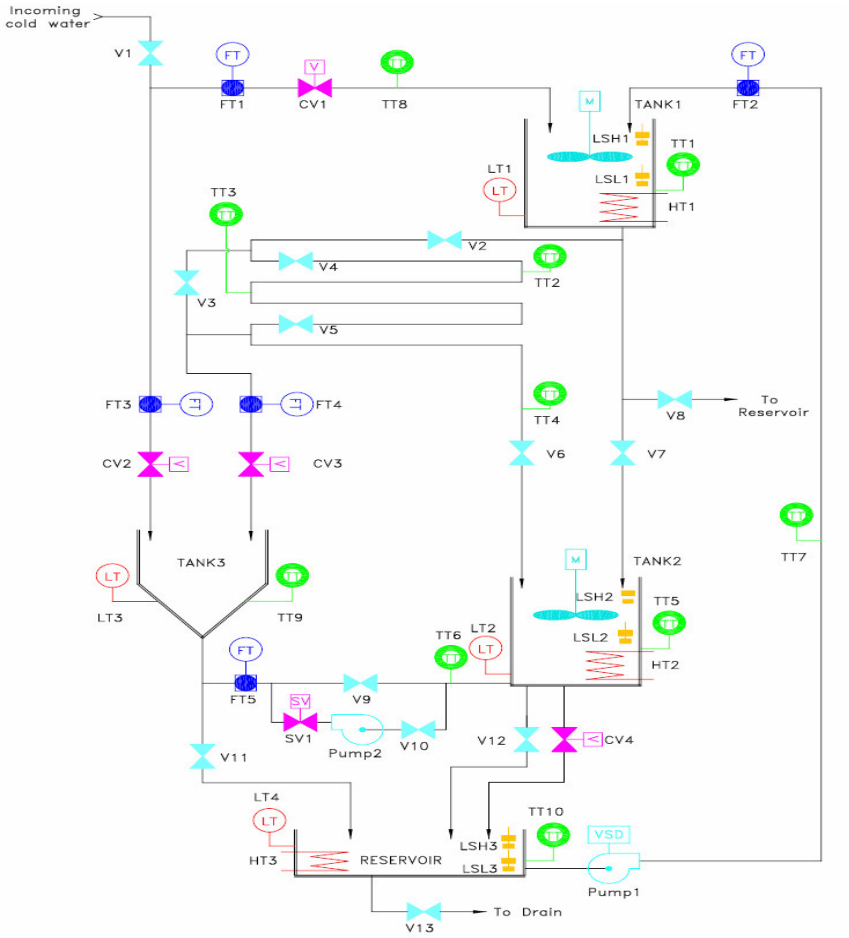

Figure 2.7: Schematic of DACS lab experimental setup

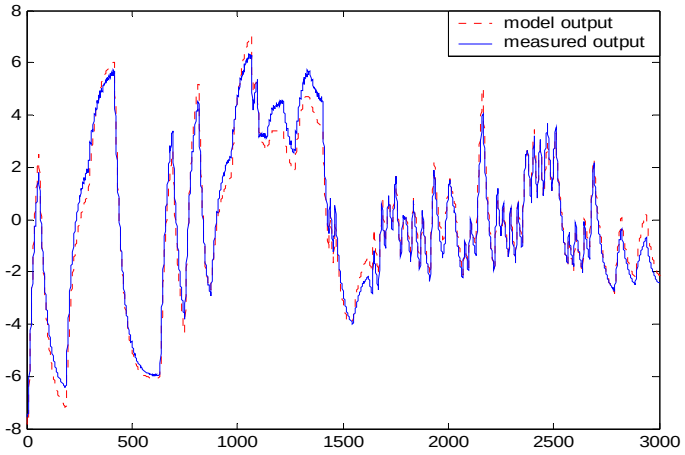

Figure 2.8: Comparison of model output and measured output data using CVA, C3

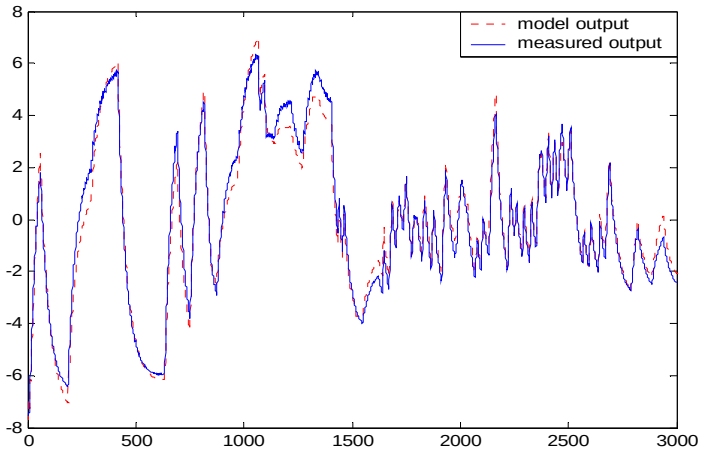

Figure 2.9: Comparison of model output and measured output data using N4SID, C3

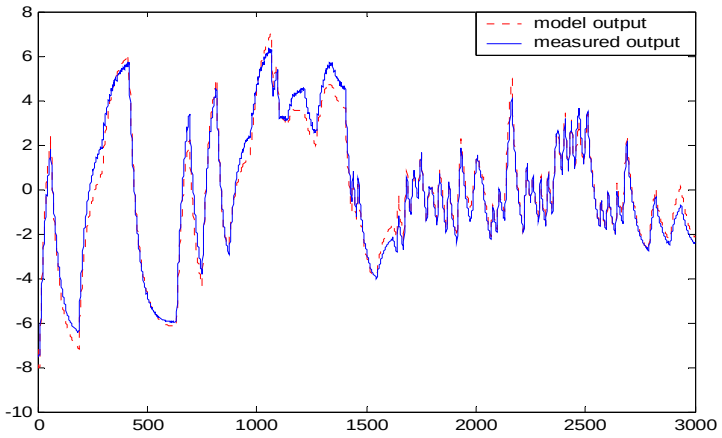

Figure 2.10: Comparison of model output and measured output data using MOESP, C3

2.5.2 Simulation Example

The pH neutralization process is very common in the many chemical and biochemical processes. First principles modeling approach gives highly nonlinear equations that involve the often unavailable equilibrium constants. A black-box modeling approach is ideal in such a scenario. In this example, we consider acid-base neutralization process performed in a single tank. The detailed system description, process model and operation conditions can be found in Henson and Seborg (1994). The level and pH of the liquid in the well stirred neutralization tank are the two outputs that are manipulated by the acid and base flow rates. In this case study, however, the system is perturbed by specially designed random buffer flow rate (shown in Figure 2.11) in which acid and base flow rates are kept constant. The pH of the neutralization tank is the output of the system. The input and output sampling interval are one second in this case. The signal to noise ratio was kept at 10 for identification purposes. The model was validated by comparing the actual and predicted output of the data obtained from a different input-output sequence. As seen in Figures 2.12, 2.13 and 2.14, the three subspace methods can identify the system quite well mainly as long as the process is around the steady state.

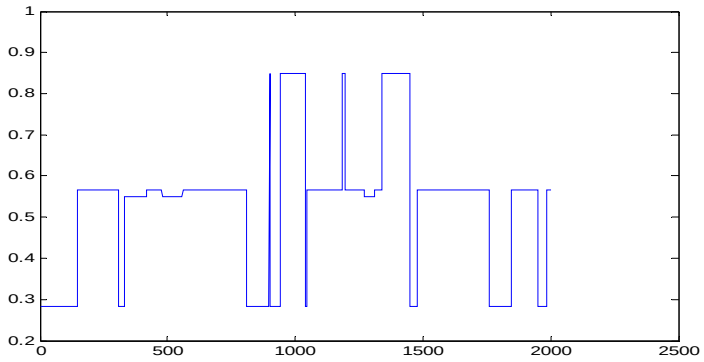

Figure 2.11: The perturbation signal (buffer flow rate) to the system

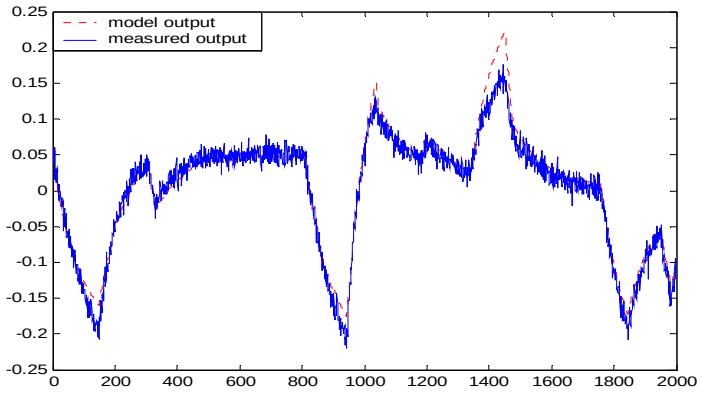

Figure 2.12: Comparison of model output and measured output data using CVA

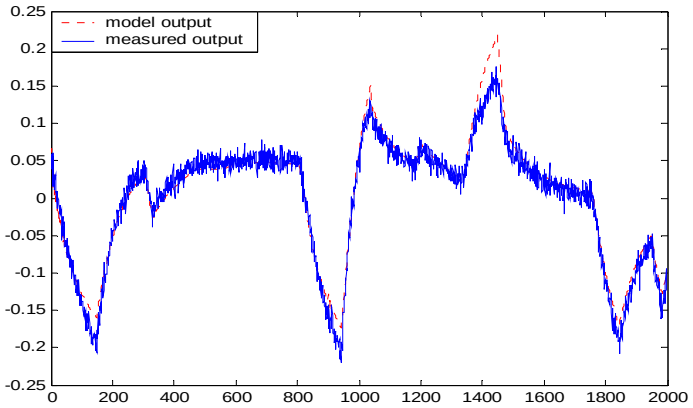

Figure 2.13: Comparison of model output and measured output data using N4SID

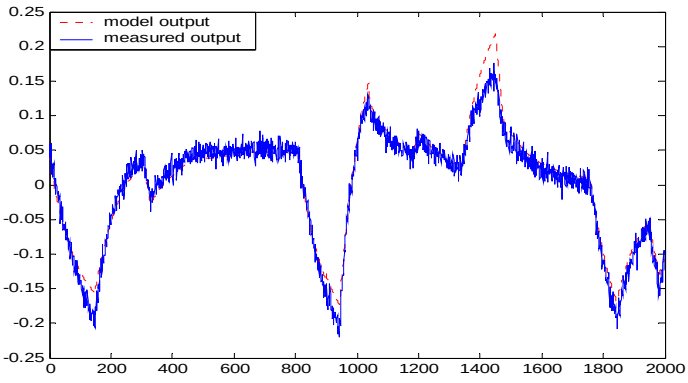

Figure 2.14: Comparison of model output and measured output data using MOESP

2.6 Conclusions

It can be concluded that the presented linear subspace identification methods: CVA, N4SID, and MOESP, are powerful tools for identification purpose even when the system shows mild nonlinearity. For retaining simplicity, we mainly illustrated the workability of these methods on single input single output (SISO) processes. In the following chapters, we will identify multiple input single output (MISO) systems with the 4SID methods.

CHAPTER 3

LIFTING

Multirate systems are periodically time varying systems and so many developed identification methods cannot be directly applied. Lifting technique is a powerful tool which converts linear periodically time varying system to linear time invariant system in which most of the system identification techniques can be applied successfully. Thus, lifting technique becomes the powerful tool in multirate system identification scenario. The availability of discrete time fast rate model is crucial in inferential control (e.g. in distillation columns, bioreactors and polymer reactors). Following the identification of the *slow rate model* using multirate data and the lifting technique, the *fast rate model* (that is useful for controller design and for output prediction) can be extracted using the method of Li et al. (2001) and Wang et al. (2004). In this chapter, we introduce the lifting technique and discuss configurations of lifted system. Application of lifting technique to multirate system identification including the extraction of the fast rate model is demonstrated. Both linear and nonlinear multirate systems are considered.

3.1 Lifting Technique and Lifted System

Kranc (1957) first introduced the lifting technique as a switch decomposition technique. Then, Friedland (1960) developed the lifting technique which converts a periodically time varying system into time invariant system in discrete domain. Further developments were made by Khargonekar et al. (1985) and his framework has since been widely adopted. Based on Li (2001) and Wang et al. (2004), the concept of lifting technique and lifted system is demonstrated in this section.

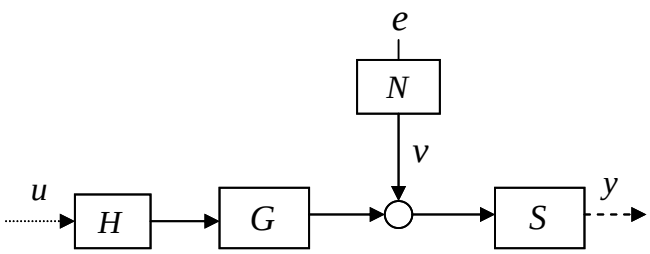

Figure 3.1: SISO multirate sampled-data system

In Figure 3.1, G is the continuous time linear time invariant (LTI) system, H and S represent the discrete time hold and sampler respectively, u and y are input and output of the process which are sampled according to H and S respectively. These assumptions hold throughout this thesis. The whole system (from u to y) is linear periodic time variant (LPTV) system. The dotted-line represents the fast rate sampling (sampling interval mp) and dash-line represents the slow rate sampling (sampling interval np), where the assumption is $m < n$ throughout this thesis (multirate systems with fast control rates and slow output sampling rates are the most common in the chemical industry), and p is the base time period. N represents the noise dynamics, e represents the noise signal, and v is the noise to the system with the fictitious sampling interval np.

For simplicity, we assume $m = 1$ in this section. The discrete time signals u_k and y_k are defined on Z_+, set of non-negative integers. The n-fold lifting operator L_n defines the mapping u to $\underline{u}$ (lifted signal):

$$\{u_0, u_1, u_2, \ldots\} \mapsto \left\{ \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{n-1} \end{bmatrix}, \begin{bmatrix} u_n \\ u_{n+1} \\ \vdots \\ u_{2n-1} \end{bmatrix}, \ldots \right\},$$

and so we can define $\underline{u} = L_n u$. It is clear that dimension of $\underline{u}$ is n times that of u and underlying period of $\underline{u}$ is n times that of u again. Thus, now $\underline{u}$ and y have the same time interval, nT and the lifted system becomes single rate system. The lifted slow rate system is linear time invariant and details can be found in Khargonekar et al. (1985). A SISO multirate system has effectively been converted into a MISO/MIMO single rate system. Standard system identification tools can now be applied to identify a model that represents the system dynamics for the slow sampling period. A fast rate model must then be extracted from this slow rate model.

The lifting operator L_n and inverse lifting operator L_n^{-1} obey the following properties:

$$L_n^{-1} L_n = I ,$$

and

$$L_n L_n^{-1} = I .$$

L_n^{-1} maps $\underline{u}$ back to u as follows:

$$\left\{ \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{n-1} \end{bmatrix}, \begin{bmatrix} u_n \\ u_{n+1} \\ \vdots \\ u_{2n-1} \end{bmatrix}, \ldots \right\} \mapsto \{u_0, u_1, u_2, \ldots\}.$$

The lifting operator also preserves norms:

$$\| Lu \|_2 = \| u \|_2 .$$

Even if the ratio of m to n is not an integer, we can apply the lifting operator and get the lifted signals.

After lifting, we get the fictitious system of Figure 3.2. Due to the above properties of lifting operators, the following Figure 3.2 representing the lifted Multirate system is identical to that in Figure 3.1.

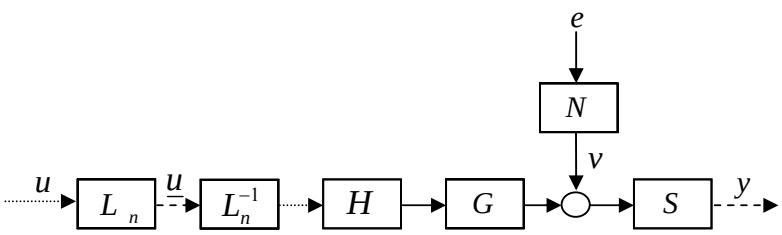

Figure 3.2: SISO lifted Multirate sampled-data system

Consider a state space model represented by the system matrices $[A, B, C, D]$.

Li (2001) expanded the state space model of the system as follows:

Let $n = \gamma$ for this case, then

$$x(\gamma k + 1) = Ax(\gamma k) + Bu(\gamma k)$$

$$x(\gamma k + 2) = A^2 x(\gamma k) + ABu(\gamma k) + Bu(\gamma k + 1)$$

$$\vdots$$

$$x(\gamma k + \gamma) = A^{\gamma} x(\gamma k) + A^{\gamma-1} Bu(\gamma k) + \ldots + Bu(\gamma k + \gamma - 1)$$

$$y(\gamma k) = Cx(\gamma k) + Du(\gamma k), \tag{3.1}$$

For the lifted system, equation (3.1) can be expressed as:

$$x_l(k+1) = A^{\gamma} x_l(k) + A^{\gamma-1} B\underline{u}_1(k) + \ldots + B\underline{u}_n(k)$$

$$y_l(k) = Cx_l(k) + \begin{bmatrix} D & 0 & \ldots & 0 \end{bmatrix} \underline{u},$$

where, $\underline{u}_1(k) = u_{nk}, \underline{u}_2(k) = u_{nk+1}, \ldots, \underline{u}_n(k) = u_{nk+n-1}$.

Then, the state space matrices of the lifted system (A_l, B_l, C_l, D_l) can be written as:

$$\left[\begin{array}{c|c} A_l & B_l \\ \hline C_l & D_l \end{array}\right] = \left[\begin{array}{c|cccc} A^{\gamma} & A^{\gamma-1}B & A^{\gamma-2}B & \cdots & B \\ \hline C & D & 0 & \cdots & 0 \end{array}\right].$$

The lifting operation causes the lifted system with increased input-output dimensions. After lifting the multirate SISO system, the system becomes MISO or multi input

multi output (MIMO) system depending on the ratio of sampling intervals. The system becomes MIMO in the case of non-integer ratio of sampling interval (e.g. $m = 2$ and $n = 3$); identification of this kind of multirate system are explained with experimental case studies in section 3.4.

Now we consider the SISO multirate system for the case where both m and n are coprime with the common base period of p. The discrete-time input signal u_k and output signal y_k are sampled at non-negative integer time set $Z_+ := \{0,1,2,\ldots\}$ with the underlying period mp and np respectively (the updating period of zero-order hold and sampler are mp and np respectively). The noise has the fictitious sampling interval same as the output. Here, we need to lift both input and output to be a single rate system with the common period mnp. The input u is lifted to $\underline{u}$ by L_n, and output y and noise v are lifted to $\underline{y}$ and $\underline{v}$ by L_m accordingly. Now the fictitious lifted system becomes as shown in Figure 3.3.

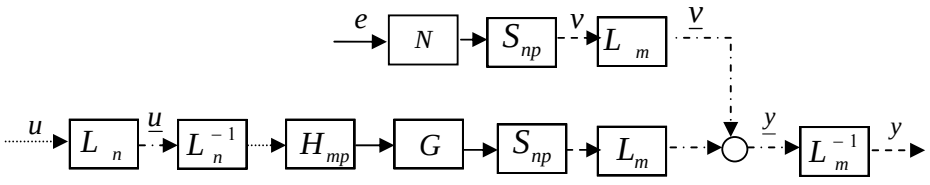

Figure 3.3: SISO lifted multirate sampled-data system when m and n are coprime

After lifting, the dimension of $\underline{u}$ becomes n times that of u and that of $\underline{y}$ becomes m times that of y, and the lifted system $\underline{G}$ becomes as follows:

$$\underline{G} = L_m S_{np} G H_{mp} L_n^{-1}. \tag{3.2}$$

In order to find the discrete-time state space model of $\underline{G}$, we discretize G via the zero order hold method to get $G_p := S_p G H_p$, where S_p and H_p are the sampler and zero-order hold with period p.

Let the state space model of the G_p is

$$D + C(zI - A)^{-1}B =: \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

and that of the $G_{mp} := S_{mp} G H_{mp}$ is

$$G_{mp}(z) = \begin{bmatrix} A_{mp} & B_{mp} \\ C & D \end{bmatrix},$$

where, $A_{mp} = A^m$, $B_{mp} = (A^{m-1} + A^{m-2} + \cdots + I)B$,

and G_{mp} is the discrete-time system with period mp.

By the identities $S_{np} = S_{np} H_p S_p$ and $H_{mp} = H_p S_p H_{mp}$, equation (3.2) becomes

$$\underline{G} = L_m S_{np} H_p (S_p G H_p) S_p H_{mp} L_n^{-1}$$

$$\underline{G} = L_m S_{np} H_p L_{mn}^{-1} L_{mn} G_p L_{mn}^{-1} L_{mn} S_p H_{mp} L_n^{-1} \qquad (3.3)$$

By these definitions:

$$\underline{S} = L_m S_{np} H_p L_{mn}^{-1}, \quad \underline{G}_p = L_{mn} G_p L_{mn}^{-1}, \quad \underline{H} = L_{mn} S_p H_{mp} L_n^{-1},$$

equation (3.3) becomes

$$\underline{G} = \underline{S}\,\underline{G}_p\,\underline{H}. \qquad (3.4)$$

A state space model of the lifted system can be expressed as follows (Khargonekar et al., 1985):

$$\underline{G}_p(z) = \left[\begin{array}{c|ccc} A^{mn} & A^{mn-1}B & A^{mn-2}B & \cdots & B \\ \hline C & D & 0 & \cdots & 0 \\ CA & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ CA^{mn-1} & CA^{mn-2}B & CA^{mn-3}B & \cdots & D \end{array}\right].$$

The constant matrices $\underline{H}$ and $\underline{S}$ are given by:

$$\underline{H} = \left[\begin{array}{cccccc} I & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ I & 0 & 0 & 0 & \cdots & \vdots \\ \hline 0 & I & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & I & 0 & 0 & \cdots & 0 \\ \hline \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \hline 0 & 0 & 0 & 0 & \cdots & I \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & I \end{array}\right]_{(mn)\times n \text{ blocks}},$$

and

$$\underline{S} = \left[\begin{array}{cccccccccccccc} I & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & I & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \underbrace{0 \quad 0 \quad \cdots \quad 0}_{n} & & & & \underbrace{0 \quad 0 \quad \cdots \quad 0}_{n} & & & & \cdots & \underbrace{I \quad 0 \quad \cdots \quad 0}_{n} & & & \end{array}\right]_{m\times(mn) \text{ blocks}},$$

where, Identity matrices I reduce to 1 if G is a SISO process.

After pre- and post- multiplying the $\underline{G}_p$ with $\underline{S}$ and $\underline{H}$, the state space model for $\underline{G}$ in equation (3.4) is:

$$\left[\begin{array}{c|cccc} A^{mn} & \sum_{\kappa=mn-m}^{mn-1} A^{\kappa}B & \sum_{\kappa=mn-2m}^{mn-m-1} A^{\kappa}B & \cdots & \sum_{\kappa=0}^{m-1} A^{\kappa}B \\ \hline C & D & 0 & \cdots & 0 \\ CA^{n} & C\sum_{\kappa=n-m}^{n-1} A^{\kappa}B & C\sum_{\kappa=0}^{n-m-1} A^{\kappa}B + D & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{mn-n} & C\sum_{\kappa=mn-n-m}^{mn-n-1} A^{\kappa}B & C\sum_{\kappa=mn-n-2m}^{mn-n-m-1} A^{\kappa}B & \cdots & 0 \end{array}\right].$$

Lifting the noise model is similar to the case in which the ratio of sampling interval is integer. Now, we get the overall lifted model for the system in which the noise is introduced as a measured disturbance:

$$\underline{y} = \underline{G}\,\underline{u} + \underline{N}\,\underline{e}\,.$$

After lifting operation, both $\underline{G}$ *and* $\underline{N}$ become LTI. Most of the statistical properties of e_k are preserved - if e_k is white noise or Gaussian, so is $\underline{e}_k$.

3.2 Identification of the Lifted Slow-rate Model

To achieve the identifiablity of a state space model, the lifted slow-rate model must be controllable and observable. The lifted slow-rate system is controllable and observable only if the continuous-time system G is observable and controllable. This assumption is valid with the non-pathological sampling interval p; the continuous time delay τ must be in the range of $[0, p]$. Wang et al. (2004) proved that the lifted system can be controllable: (A_l, B_l) is controllable if (A, B) is controllable and A has no eigenvalue on the unit circle (the proof can be seen in Wang et al. (2004)). If the continuous time delay τ is larger than the base sampling period p, the system loses observability. It is impossible to extract fast rate model from the lifted model if the lifted model loses observability. The remedy proposed by Li et al. (2001) is as follows:

Step (1): Finding the time delay matrix Ψ by applying the standard correlation analysis to lifted signals $\underline{u}$ and $\underline{y}$.

Let

$$\Psi = \begin{bmatrix} l_{00} & l_{01} & \cdots & l_{0,n-1} \\ l_{10} & l_{11} & \cdots & l_{1,n-1} \\ \vdots & \vdots & \cdots & \vdots \\ l_{m-1,0} & l_{m-1,1} & \cdots & l_{m-1,n-1} \end{bmatrix},$$

where l_{ij} = estimated time delay (non negative integer) from $\underline{u}_j$ to $\underline{y}_i$,

$$\underline{u}_j = j^{\text{th}} \text{ lifted input signal } (j = 0,1,\ldots,n-1),$$

$$\underline{y}_i = i^{\text{th}} \text{ lifted output signal } (i = 0,1,\ldots,m-1).$$

Taking the lifting effect into account, the actual time delay τ_{ij} from $\underline{u}_j$ to $\underline{y}_i$ be:

$$\tau_{ij} = \tau + jmp - inp,$$

where, τ = continuous time delay.

Meanwhile, τ_{ij} is in the following interval:

$$(l_{ij} - 1)mnp < \tau_{ij} \le l_{ij} mnp.$$

So, the relation between l_{ij} and τ is:

$$(l_{ij} - 1)mnp < \tau + jmp - inp \le l_{ij} mnp.$$

Step (2): Estimating the time delay $\hat{\tau}$ that has one to one correspondence between Ψ and positive integer k.

$$kp < \hat{\tau} \le kp + p.$$

Step (3): Finding the shifting operator κ_1 and κ_2.

Since m and n are coprime,

$$k = \kappa_1 m + \kappa_2 n.$$

Step (4): Shifting the measured input to right by κ_1 and measured output to left by κ_2.

By shifting the input and output, we can maintain the time delay between shifted signals is not larger than p, so that we can maintain the observability and controllability of the lifted system also.

There exits a causality constraint in the lifted model. Li et al. (2001) proposed a modified subspace identification algorithm to deal with such a constraint. Wang et al.

(2004) proposed a structured state space model with free parameters as an easier alternative.

3.3 Computing the Fast-rate Model

There are three ways (Li et al. (2001)) to extract the fast rate model from lifted slow-rate system. Wang et al. (2004) further developed these methods and demonstrated getting a fast-rate model with sampling period p for the system with $3p$ hold interval and $2p$ sampling interval. In this section, we present three methods to get a fast-rate model with the sampling period p for the mp hold interval and np sampling interval, in which $m < n$ and both are prime numbers.

3.3.1 Matrix Roots Approach

This method is derived from the following identity of the lifted model:

$$A_l = A^{mn}. \tag{3.5}$$

Let the pole of the continuous process G be

$$\eta_l = \alpha_l + i\beta_l,$$

and the corresponding poles of discretized system with interval mnp is

$$\sigma_l = e^{mnp\eta_l}. \tag{3.6}$$

Then the equation (3.6) becomes,

$$\sigma_l = e^{mnp\alpha_l} e^{imnp\beta_l}. \tag{3.7}$$

Under the assumption $|mnp\beta_l| \le \pi$, we can get the matrix A of the fast-rate model with underlying period p as $A = A_l^{1/mn}$.

3.3.2 Eigenvalue Approach

Eigenvalue approach is based on the condition that A is diagonalizable:

$$W^{-1}A_l W = diag\{\lambda_1, \lambda_2, \ldots, \lambda_\omega\},$$

where,

λ_l, $(l = 1,2,\ldots,\omega)$ is eigenvalues of A_l, and W is corresponding eigenvector matrix.

By equation (3.5), A and A_l share the same eigenvectors.

Now, each eigenvalues are again

$$\lambda_l = e^{mnp\eta_l} = e^{mnp\alpha_l} e^{imnp\beta_l}.$$

Via the assumption $|mnp\beta_l| \le \pi$,

$$A = W\ diag\{\lambda_1^{\frac{1}{mn}}, \lambda_2^{\frac{1}{mn}}, \cdots, \lambda_\omega^{\frac{1}{mn}}\}\ W^{-1}.$$

The finding of other three matrices i.e. matrices B, C and D is the same for these two methods.

$$C = C_1$$

$$B = \left(\sum_{i=0}^{m-1} A^i\right) B_n$$

B_n and C_1 can be obtained from the lifted model as:

$$B_l = \begin{bmatrix} B_1 & B_2 & \cdots & B_n \end{bmatrix},$$

and

$$C_l = \begin{bmatrix} C_1^T & C_2^T & \cdots & C_m^T \end{bmatrix}^T,$$

where,

B_i $(i = 1,2,\ldots,n)$ is $q \times 1$ column vector,

C_j $(j = 1,2,\ldots,m)$ is $1 \times q$ row vector.

3.3.3 Alternate Approach

Though theoretically sound, the above two methods sometimes present numerical difficulties. An alternate approach is proposed here as a practical solution to the problem. Firstly, we employ model reduction to the slow-rate model to obtain minimal state space form. The reduced-order model is produced with matching DC gain using equivalent steady state step response. The state or states to be deleted is determined using 'balreal' command in Matlab. The 'balreal' command (The Math Works, Inc. 1998) is used for producing a balanced realization in state space form reflecting the same controllable and observable properties of the individual states. The elements in the diagonal of the balanced realization form reflect the grammian-based combined controllable and observable properties of the different states. We can delete those elements of the diagonal (states) with small value so that the most important feature of the original system can be captured by retaining the larger values of the diagonal elements. We deleted the weak state or states which are computed from 'balreal' command (by deleting the small values of the diagonal elements). After deleting the weak state or states using the Matlab command 'modred', the remaining model contains the most essential input-output character of the original slow-rate system.

The 'modred' command (The Math Works, Inc. 1998) with matching DC gain method works as follows for the discrete-time state space model:

Let the discrete-time state space model be

$$x(k+1) = Ax(k) + Bu(k)$$

$$y(k) = Cx(k) + Du(k).$$

The state vector is divided into two parts, x_1 and x_2. x_1 are the states to be retained and x_2 are the states that may be eliminated.

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(k)$$

$$y(k) = \begin{bmatrix} C_1 & C_2 \end{bmatrix} x(k) + D\, u(k)$$

Then x_1 states are calculated by setting the derivative of x_2 to zero, and the reduced-order model is as follows:

$$x_1(k+1) = [A_{11} - A_{12}A_{22}^{-1}A_{21}]\, x_1(k) + [B_1 - A_{12}A_{22}^{-1}B_2]\, u(k)$$

$$y(k) = [C_1 - C_2A_{22}^{-1}A_{21}]\, x(k) + [D - C_2A_{22}^{-1}B_2]\, u(k)$$

Then the fast rate model with p sampling interval is extracted from resulting low order slow-rate discrete-time model using 'd2d' Matlab command (this command can transform discrete-time model with particular sampling interval into discrete-time model with required sampling interval). This method operates in state space domain and resulting fast-rate model is also in discrete-time state space form.

3.4 Linear System Identification

For the identification of linear systems, we use the above mentioned algorithms in a straightforward manner. The procedures and application of the algorithm for the linear system identification of multirate system is presented with the example in this section. The data used in this example are obtained from experimental setup in the computer process control laboratory at the University of Alberta. The equipment considered is a pilot scale stirred tank heater. The configuration of the process is the same as the Case Study I of Chapter 2 and a schematic of this process can be found in Li et al. (2001). In this example, the manipulated input (cold water valve position) and the measured

output (tank water level) were sampled at every 80 sec and 120 sec intervals respectively. Thus, this system became SISO multirate system with $m = 2,\ n = 3$, and base period $p = 40$ sec. At the same time, the input and output were sampled at every 40 sec to obtain a fast-rate data for validation purpose. The details of the process conditions and input-output configuration can be found in Wang et al. (2004).

The ratio between m and n is rational number and this is the general multirate system. The identification procedure for this multirate system is as follows:

Step (1): Lifting the input-output multirate signal to be a single-rate slow system.

We lift input signal by L_3 into $\underline{u}$ and output signal by L_2 into $\underline{y}$. Now $\underline{u}$ has three inputs and $\underline{y}$ have two outputs (MIMO) with common sampling interval $mnp = 240$ sec, and the multirate system (linear periodically time varying (LPTV) system) gets transformed into a LTI system.

Step (2): Estimate the time delay $d_c p$ of the continuous-time process.

The continuous time delay $d_c p$ is estimated based on lifted signal $\underline{u}$ and $\underline{y}$, and the measured signals are shifted if there exits time-delay applying the Li et al. (2001) algorithm mentioned in section 3.2.

Step (3): Identification of slow-rate system.

The slow-rate system is identified using any of the subspace based state space identification methods (i.e. N4SID, CVA or MOESP) to the shifted data.

Step (4): Extracting fast-rate model.

The fast-rate model with sampling period 40 sec is extracted from identified slow-rate model using any of the methods mentioned in section 3.3 and the method described in section 3.5.1.

Step (5): Associating the time delay.

Associate the estimated time delay from step (2) (in discrete-time domain) to the resulting fast-rate model.

These steps were applied in sequence to the multirate data that was collected. The slow rate model was estimated using the CVA method (N4SID method gives an unstable model for this case). The fast rate model was extracted from this slow rate model using the modified alternate approach (modified model reduction approach) proposed in this thesis (see section 3.5.1). After lifting the system, the lifted slow-rate system becomes MIMO single rate system in this case. Here is the little modification for the modified alternate approach. This modification involves choosing the proper lifted I/O pair that can give the maximum possible system information. This lifted I/O pair is chosen based on the mean square error (mse) between available measured slow sampled data and estimated slow-rate model output.

The fast rate model identified from multirate data is given by $G_{MR}(s)$ and in transfer function form it is equal to $\frac{0.006683z^2 + 0.2138z + 0.1872}{z^2 - 1.476z + 0.5211}$. The model identified from single rate data (i.e. input and output both sampled at 40 sec) is $G_{SR}(s)$ and it is equal to $\frac{-0.004291z^2 + 0.005784z + 0.8902}{z^2 - 0.9609z + 0.0595}$. The step response models obtained from these different identified models are very close to each other and can be seen in Figure 3.5. The fast-rate model identified from multirate data is validated by predicting model outputs to a specific input sequence. The comparison of the estimated fast-rate model prediction to the actual measured output is shown in Figure 3.4. The results indicate that the model identified from multirate data is pretty good.

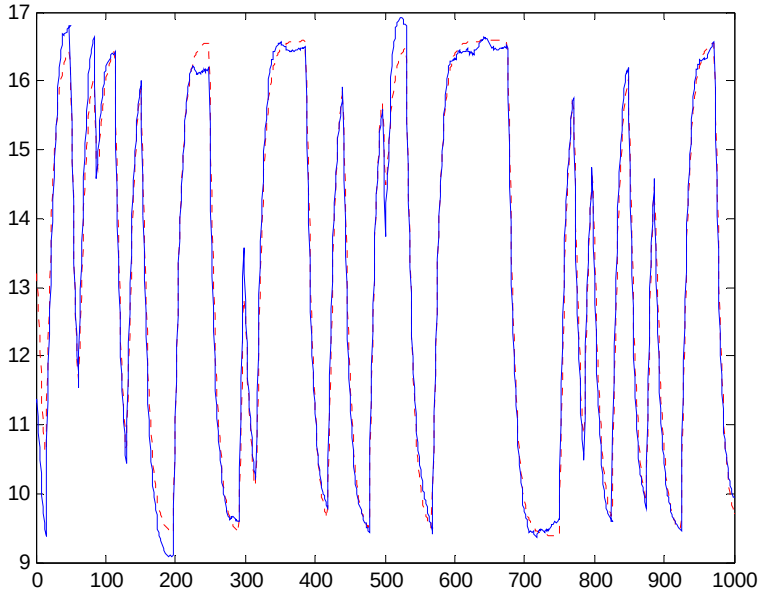

Figure 3.4: Comparison of estimated fast-rate model output (dashed line) and measured output (solid line) using modified alternate approach

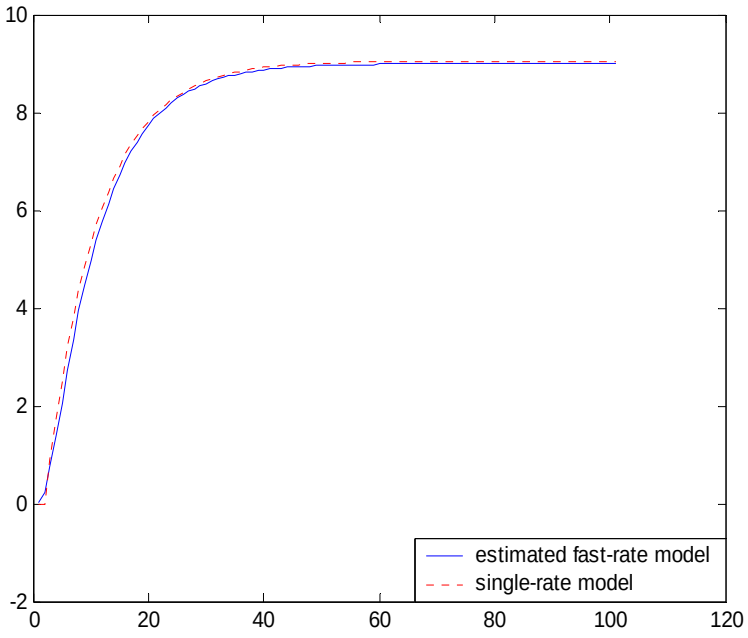

Figure 3.5: Comparison of step response models obtained from estimated fast-rate model and single-rate model

3.5 Nonlinear System Identification

Most of the processes are nonlinear in nature. Among the various nonlinear models, Hammerstein and Wiener models are useful representations for chemical processes. For these reasons, it would be appropriate to develop multirate system identification methods for Hammerstein and Wiener models. To identify the Hammerstein and Wiener models, we use separable nonlinear least squares (SLS) method to estimate the parameters of the linear dynamic and nonlinear static polynomial. The parameters of the linear dynamic part of Wiener model is estimated as demonstrated in Bruls et al. (1999). The initial estimate of linear dynamic part is identified in subspace domain using methods such as CVA or MOESP.

From our research, it is found that different γ values (ratio of sampling intervals) affect the gain of estimated fast-rate model. It is also found that the identification of slow-rate model sometimes will not give the exact estimation of gain of the true model. These factors point out a need to adjust the gain of the identified slow-rate model since the identified slow-rate model is the only source from which the fast-rate model can be obtained. A new procedure is developed by modifying the "alternate approach" proposed for linear systems. This "modified alternate approach" is presented below.

3.5.1 Modified Alternate Approach

The modified alternate approach is that in which the identified fast-rate models are estimated by adjusting the gain of the identified slow-rate model. The aim of adjusting the gain of the identified model is to obtain as much as the same magnitude of the

gain of the true model. The characteristics and procedure of adjusting the gain of identified slow-rate model is as follow:

(1) The value to adjust the gain called 'adj_gain' is the value to multiply the B matrix of identified slow-rate state space model.

(2) adj_gain is estimated based on the available output data of multirate data set.

(3) This method is only applicable to the alternate approach (model reduction using the step response of identified slow-rate model).

(4) Firstly, the proper lifted input signal is chosen which has the least mean square error (mse); this mse is calculated by squaring the difference between slow-rate model output and measured multirate output data. This mse is found for every lifted input signal and the lifted input signal which has least mse is chosen as "optimal" input signal.

(5) Secondly, the adj_gain is estimated by minimizing the mse of the difference between the slow-rate model output and measured output data. Multidimensional unconstrained nonlinear minimization (Nelder-Mead) routine and simplex (direct search) method are employed to find the optimum adj_gain value.

(6) After knowing adj_gain, the B matrix is multiplied with adj_gain, and the resultant matrix is called new B matrix.

(7) Finally the fast-rate model is extracted from identified slow-rate state space model using the new B matrix in place of the originally estimated B matrix.

3.5.2 Multirate Hammerstein Model Identification

The Hammerstein model identification procedure for the single-rate system was first developed by Narendra and Gallman (1966). They used an iterative method and estimated the dynamic linear subsystem and static nonlinear subsystem alternately.

This was followed by the one-step non-iterative method in which the static nonlinearity (generally expressed as a polynomial) was expanded into a series and these expansion terms were used as inputs to the linear dynamic system. This approach transformed the SISO nonlinear system into a multi-input linear time-invariant system. The standard linear identification methods can be applied for identification of this type (Stoica and Söderstörm(1982)). After that, a two-step non-iterative method was developed by Pawlak (1991); firstly the linear dynamic part was identified. In the second step, this information was used in identifying the nonlinearity. Westwick and Kearney (2001) explored a technique to identify Hammerstein model using SLS in which polynomial with predefined order is used to estimate the static nonlinearity and impulse response function (using a correlation analysis) is used to estimate the linear dynamic part. They used iterative nonlinear optimization routine (Levenberg-Marquardt iterations) and their work was based on the Hunter and Korenberg (1986) iterative algorithm. In this work, a new algorithm for the identification of multirate Hammerstein type nonlinear system with the assumption that the nonlinear map $\Phi(.)$ can be represented by a polynomial surface of fixed pre-specified order is proposed. The method uses 4SID techniques for estimating the linear dynamic part of the model. The univariate polynomials are defined to estimate the nonlinear static part in Hammerstein model identification in this algorithm.

We consider the nonlinear system as follows:

$$\Phi(\chi) = a_1 + a_2\chi + a_3\chi^2 + \ldots + a_n\chi^f,$$

$$\mu(k) = \Phi(u(k))$$

$$x(k+1) = Ax(k) + B\mu(k)$$

$$y(k) = Cx(k) + D\mu(k) + v(k)$$

where f is the order of the polynomial and $a_i (i = 1, \ldots, n)$ are the polynomial coefficients.

The crucial requirement for the identification of this type of nonlinear system is that the input signal or signals must be persistently exciting for the data set of finite length. The algorithm for the identification of multirate Hammerstein model using lifting technique is as follows:

(1) The optimum model order is chosen using the linear identification method assuming the system is linear (this assumption provides us with the approximation of the order of the linear model) for MOESP framework, but the lifted input and output signals are employed (we assume that input is fast sampling and output is slow sampling) since we are dealing with multirate system and lifting technique.

(2) The parameters of the zero-memory gain (nonlinearity) are estimated using given order polynomials (the fast sampling input signal is used here) by iterative search (the initial estimate for static nonlinearity is provided). The objective function here is to minimize the mean square error between the estimated model output and measured output. The details of this step are:

The estimated intermediate model output that comes out from the nonlinear part is lifted using the lifting operator with respect to γ (the ratio of sampling interval of output to input). The state-space quadruple matrices of slow-rate system for linear dynamic parts are estimated (with pre-estimated model order) using the lifted signals and then the estimated state space matrices are used to update the nonlinear part. The whole model output is estimated using estimated polynomial and estimated state space matrices. Then mean square error between estimated model output and measured output is calculated for each iteration step (this is also objective function). SLS is implemented to minimize the number of parameters to be estimated (the

parameters of linear dynamic part implicitly affects the estimation of static nonlinear part (parameters of polynomial) and Gauss-Newton optimization routine is also used.

(4) Then, the parameters for linear dynamic part are determined using the estimated intermediate model output of the nonlinear part as the input (that is lifted according to γ value) using linear MOESP identification method (CVA can also be implemented instead of MOESP).

(5) Finally, the fast-rate model is extracted from identified slow-rate model using matrix roots approach or alternate approach (including modified alternate approach). There is no need to find the fast-rate model for static nonlinearity since we identified the static nonlinearity with the fast sampling input signals.

3.5.2.1 Application with experimental data set

The usefulness of the above developed algorithm will be illustrated with the heat exchanger data set from Eskinat et al. (1991). This data set was used as Case Study II in the previous chapter (section 2.5.1.2) to demonstrate the linear subspace identification methods. It was pointed out that there is some nonlinearity associated with this system (and the data set). In this section, multirate data sets are constructed for different γ (γ = 1 to 5) and the new algorithm is tested by employing it for the identification of a Hammerstein model. The modified alternate approach is used in the extraction of fast-rate models from different identified slow-rate models. The matrix roots approach was also tried to arrive at the fast rate models but it was found that the modified alternate approach gave a much better model fit. Table 3.1 provides a comparison of mean square error (mse) of the modified alternate approach (MAA) and the matrix roots approach (MRA). The mse is obtained by squaring the difference

between the fast-rate model predicted output and the measured fast-rate output data. The superiority of the modified alternate approach developed in this thesis is evident. The improvement is mainly due to the estimation and correction for the process gain.

Table 3.1. Mean Square Error Comparison

γ	**MRA**	**MAA**
1	27.818	27.773
2	81.077	21.976
3	108.38	32.68
4	109.59	62.271
5	2033.7	326.74

These results are obtained using the heat exchanger data set with the proposed multirate Hammerstein model identification method. The linear dynamic model component is assumed to be third order and nonlinear static polynomial is taken to be a fourth order polynomial. The estimated fast-rate model outputs obtained from MAA are cross validated with measured fast rate data set for each γ in Figures 3.6 through 3.10. As seen from these Figures and also from Table 3.1, the model quality deteriorates (as expected) as γ increases. This also points out that sophisticated identification approaches can do only so much; if the data is not sampled well or is not of good quality, the identification results will, in general, be poor.

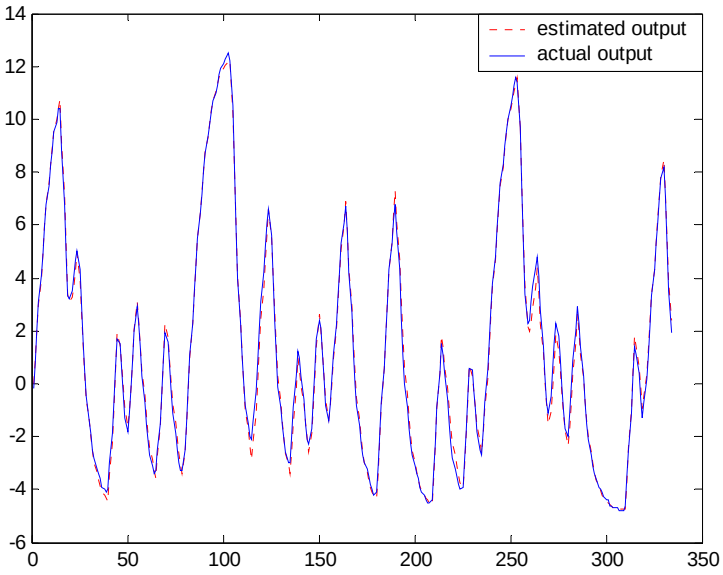

Figure 3.6: Cross validation for $\gamma = 1$

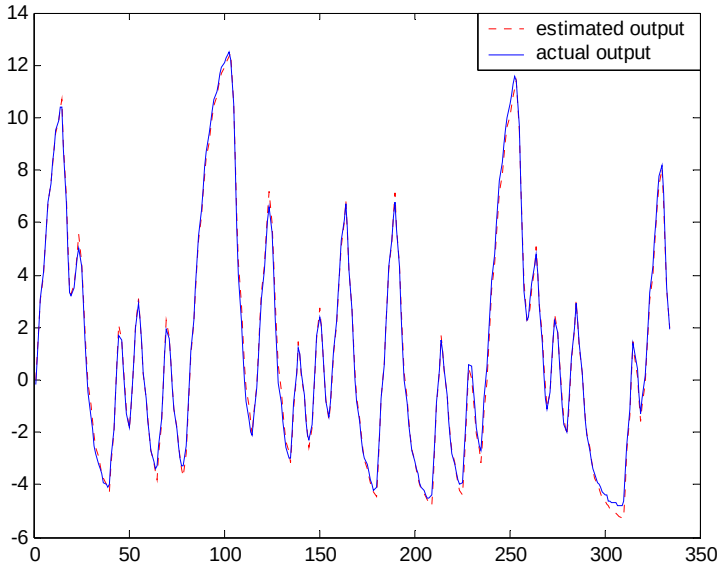

Figure 3.7: Cross validation for $\gamma = 2$

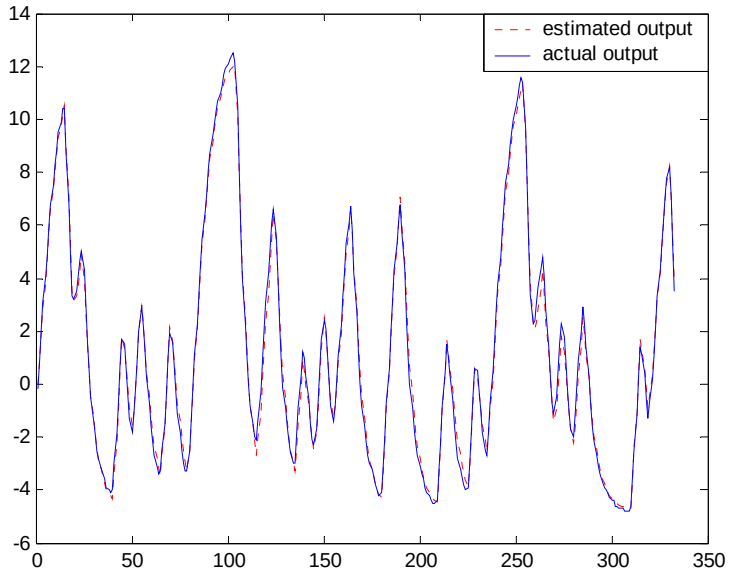

Figure 3.8: Cross validation for $\gamma = 3$

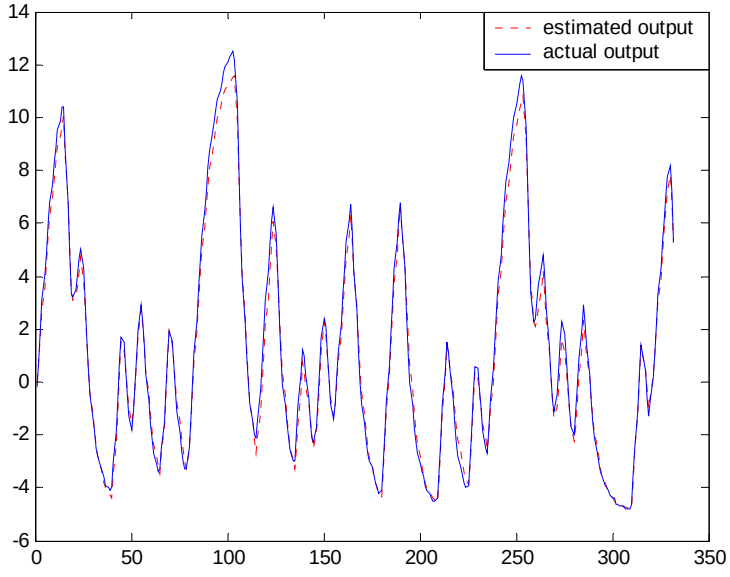

Figure 3.9: Cross validation for $\gamma = 4$

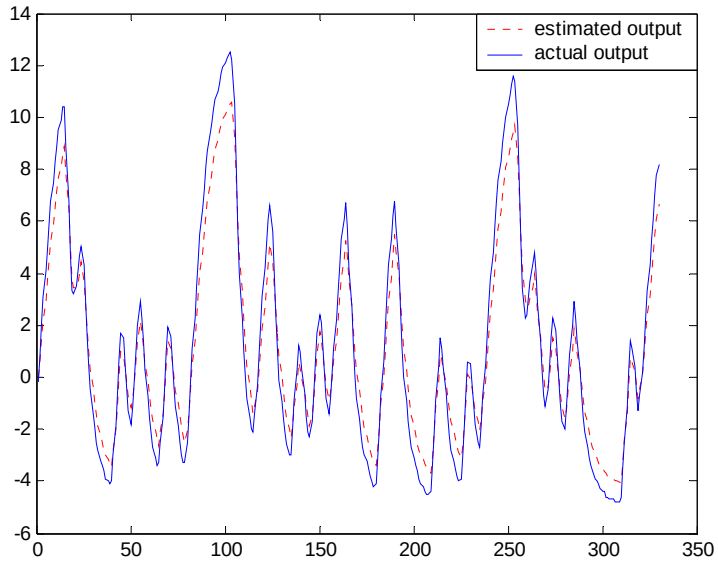

Figure 3.10: Cross validation for $\gamma = 5$

3.5.3 Multirate Wiener Model Identification

Wiener model identification is of interest to many researchers as it can represent processes such as pH neutralization, distillation columns and polymer reactors. Boyd and Chua (1985) showed that this type of model can model time invariant systems with fading memory. Bruls et al. (1999) developed two algorithms for the identification of this type of nonlinear system; the linear part was identified as a state space model and the parameters of nonlinear static part were estimated by a linear combination of basis functions (Tchebyshev polynomials) in algorithm W1. The LTI part of the Wiener system is the only unknown and static nonlinear part is known (e.g. a sensor with saturation) in algorithm W2. All previous work regarding this type of nonlinear identification was done for single-rate system only. In our work, multirate nonlinear identification for Wiener model is developed based upon the work of Bruls et al. (1999). Again, SLS is used to obtain the parameters of the linear dynamic part

and the nonlinear memoryless part (with predefined maximum order of polynomials and state space model order). We use the lifting technique to identify the slow rate system of linear dynamic part using the MOESP algorithms developed by Verhagen coworkers (Verhagen and Dewilde (1992a, 1992b), Verhagen (1993), and Verhagen (1994)).

Tchebyshev Polynomials

$$\phi(\chi) = \varphi_0 T_0(\chi) + \ldots + \varphi_n T_n(\chi)$$

where, $T_i (i = 0, \ldots, n)$ is i^{th} order Tchebyshev polynomial.

Tchebyshev polynomials are often chosen by researchers for the identification of processes and generally in cases where model approximations are needed. Tchebyshev polynomials are a ratio of polynomials, rational functions and more accurate estimates can be obtained by using this kind of polynomials in computational work. The other reason is that Wiener model identification is more challenging as compared to the identification of Hammerstein models; in this regard, Tchebyshev polynomials might be a better choice than the polynomial functions employed in Hammerstein model identification.

Algorithm

The nonlinear system of Wiener type as in this mathematical form is considered for identification of this algorithm:

$$x(k+1) = Ax(k) + Bu(k)$$

$$y(k) = Cx(k) + Du(k)$$

$$\vartheta(k) = \phi(y(k)),$$

where $\vartheta(k)$ is the output of the system.

The algorithm for the identification of Wiener type nonlinear multirate system can be summarized as follows:

(1) The model order of the linear dynamic part is estimated with the PO algorithm of MOESP family (Verhagen (1994)) using lifted input-output data (the assumption of fast input sampling and slow output sampling is made).

(2) The initial estimate of state space quadruple matrices are determined using the subspace algorithm developed by Westwick and Verhaegen (1996) in which the PI scheme of MOESP family (Verhagen (1993)) is used. The simulated output from this initial estimate of the linear system is used to obtain the initial parameter estimates for the nonlinear static part of the model (using Tchebyshev polynomials).

(3) The parameters of the linear dynamic subsystem are calculated using the algorithm W2 (Bruls et al. (1999)) in which the Gauss-Newton optimization routine and the separable least squares (SLS) technique are used. The parameters of (A, C) pair (that is nonlinear part in SLS problem) is estimated using the output normal form in which observability grammian is transformed into identity.

(4) The slow-rate output from the state space linear dynamic model resulting from the above optimization routine and measured slow-sampled output data are employed for the calculation of the final estimate of parameters of nonlinear static subsystem.

(5) The model fit for the whole system is measured with the variance-accounted-for (VAF) metric which is defined as

$$\text{VAF}=\left(1-\frac{\text{var}(\hat{\vartheta}(k)-\vartheta(k))}{\text{var}(\vartheta(k))}\right)\times 100\%,$$

where, $\hat{\vartheta}(k)$ is the estimated output of the model, and

var (.) stands for the variance of an arbitrary sequence.

(6) The fast-rate model of the linear dynamic model is extracted from the identified slow-rate model using the matrix roots approach or alternate approach.

Case studies involving the identification of Wiener models from multirate data will be provided in Chapter 5.

3.6 Conclusions

In this chapter, a review of the lifting technique and previous developments regarding the extraction methods of fast-rate model have been provided. Two new methods (viz. alternate approach and modified alternate approach) are proposed supported by case studies. Furthermore, nonlinear multirate identification algorithms (which are developed from previously developed methods for nonlinear single-rate system) for well known Hammerstein and Wiener models are presented. The numerically stable SLS technique is used to reduce the number of parameters for both models. The developed multirate Hammerstein model identification algorithm and the effectiveness of modified alternate approach were simultaneously evaluated with an example.

CHAPTER 4

DATA SELECTION AND REGRESSION METHOD

Process models in the chemical industry usually involve several input variables. Also, in most applications, the sampling interval ratio between the output and the inputs is high (typically more than 15). Multirate identification based on lifting is considered to be ineffective under such circumstances owing to the explosive increase in the number of input variables. Alternate approaches are definitely needed. Lakshminarayanan (2000) developed a method called Data Selection and Regression (DSAR) for the identification of multirate systems. In the DSAR approach, a standard regression model is constructed for the Multirate system and the impulse response coefficients of the multirate system are estimated. The estimated impulse response coefficients may then be transformed into other forms (e.g. step response coefficients for use in MPC schemes) for controller design purposes. While the lifting technique needs regularly sampled input/output (I/O) data, DSAR can handle irregularly sampled I/O data as well. The other advantage of DSAR is that it can be applied for large (output to input) sampling interval ratios. Concepts such as optimal window size and optimal lag combination are used in order to minimize the mean square error to obtain a parsimonious regression model. Ordinary least squares (OLS), principal component regression (PCR) or partial least square (PLS) may be used to solve the regression problem. PCR and PLS are methods that can handle regression problems with highly correlated data sets. MacGregor et al. (1991) also investigated the use of PLS to regression problems involving correlated data sets but their work was concerned with single-rate system.

4.1 DSAR

The concept of DSAR is based on the well-known Finite Impulse Response (FIR) model of single-rate system. DSAR method also shares the advantages of FIR model identification such as its ability to model any complex dynamical system. DSAR overcomes the disadvantage of needing a long model kernel of FIR model by using the concept of optimal window size. DSAR can also be easily extended to model nonlinear systems by including nonlinear variables (e.g. quadratic and interaction terms) in the regressor. The basic DSAR method is explained below.

Multiple Linear Regression

DSAR employs the multiple linear regression (MLR) method for the model building purposes. MLR is the regression model in which the response variable is a function of one or more predictor variables. The response variable is "fitted" by linear combination of predictor variables. Using DSAR, we pick up the sample for which the output measurements are available. The output variable corresponding to these sampling instants are stacked into column vector Y.

For each of these sampling instants (j), a row vector x is created as follow:

$$x_j = \left[u_1(j-1) \ldots u_1(j-m_1) \ldots u_r(j-1) \ldots u_r(j-m_r)\right]$$

where, r is number of process inputs, and

$m_1 \ldots m_r$ stand for expected time to steady state for each of the inputs.

Then the row vectors, x_j's are stacked into a matrix of inputs X. Thus, it simply becomes the MLR model in which the output of the system (constructed column vector Y) is the response variable and inputs of the system (constructed matrix X) are the predictor variables.

The standard regression model for $j+N$ observation is obtained in matrix form as

$$\begin{bmatrix} y_j \\ y_{j+1} \\ \vdots \\ y_{j+N} \end{bmatrix} = \begin{bmatrix} u_{1\,j-1} & \cdots & u_{1\,j-m_1} & \cdots & u_{r\,j-1} & \cdots & u_{r\,j-m_r} \\ u_{1\,j} & \cdots & u_{1\,j+1-m_1} & \cdots & u_{r\,j} & \cdots & u_{r\,j+1-m_r} \\ \vdots & & \vdots & & & \vdots & \vdots \\ u_{1\,j+N-1} & \cdots & u_{1\,j+N-m_1} & \cdots & u_{r\,j+N-1} & \cdots & u_{r\,j+N-m_r} \end{bmatrix} \begin{bmatrix} b_j \\ b_{j+1} \\ \vdots \\ b_{j+N} \end{bmatrix} + \begin{bmatrix} e_j \\ e_{j+1} \\ \vdots \\ e_{j+N} \end{bmatrix}, \quad (4.1)$$

where $b_j \cdots b_{j+N}$ are impulse response coefficients of the system, and

$e_j \cdots e_{j+N}$ are the errors associated with each observation.

We can extend the DSAR method to MIMO systems by placing other output variables (constructed in similar way of column vector Y) besides the column of the first output variable (stacking them side by side). The noise or error matrix would be of the same size as the matrix on the left hand side of the regression equation.

4.2 Methods for Solving DSAR

4.2.1 DSAR Identification Using Ordinary Least Squares (OLS)

DSAR model, the standard regression model obtained in equation (4.1) can be expressed in standard form as

$$Y = XB + E\,.$$

The B matrix which contains impulse response coefficients is calculated based on the least squares error which is as

$$e(t) = y(t) - \hat{y}(t)\ .$$

The standard least squares problem,

$$\min_b (Y - XB)^T (Y - XB)$$

can be solved as

$$B = \left[X^T X\right]^{-1} X^T Y\,. \quad (4.2)$$

The least-squares estimator for FIR model gives consistent estimate when the number of observations tends to infinity, and is statically unbiased - the expectation of the estimate equals the true value and if the disturbance (error term) is independent of the input. The proofs of these statements can be found in the literature (e.g. Söderström and Stoica (1989), Zhu (2001)). Thus, DSAR model estimation using least-squares estimator is also consistent and statically unbiased if the mentioned assumptions are satisfied.

4.2.2 DSAR Identification Using PCR and PLS

It can be said that data are correlated when there is linear association among the data; the values of variables tend to increase or decrease together. Correlation is measured by a correlation coefficient (e.g. Pearson correlation coefficient). When the correlation exists between the variables of X , its inversion becomes problematic and the parameters cannot be determined. This is the ill-conditioned problem. These problems can be overcome by using PCR or PLS. PCR is based on the principal component analysis (PCA). The scores and loadings matrices of PCR are calculated by using PCA. PCR solves the collinearity problem by replacing the original X variables with the new basis space – a set of latent variables that are orthogonal and can span the multidimensional space of X . The redundancies in the X block are eliminated by PCR and in this process significant reduction in noise is also achieved. However, with PCR, there is always the chance that some information about the system may be lost in the discarded components. PLS is yet another alternative to MLR. Without getting into the details of the PLS algorithm (which can be obtained from Kresta et al., 1991), it suffices to say that the PLS technique generates latent variables that are more predictive of the Y variables. Both PCR and PLS are therefore

capable of circumventing the collinearity problem and produce reasonable (though biased) parameter estimates.

4.2.3 Fast-rate Step Response Model

The convolution models, impulse response model and step response model are obtained by using DSAR and they can represent free responses only. These convolution models are very intuitive to operators and plant personnel in spite of being less compact as compared to the transfer function or state space models. Moreover, these models can easily represent complex dynamics and are easy to develop from plant tests. The step response model can be obtained by processing the impulse response coefficients contained in the B matrix which is obtained by using OLS or PCR or PLS. The step response coefficients can be calculated from obtained impulse response coefficients as follows:

Let us define the impulse response model and the step response model as

$$y(t_k) = \sum_{i=0}^{k} h(i)\, u(k-i)\,,$$

$$y(t_k) = \sum_{i=0}^{k} s(i)\, u(k-i)$$

respectively. Here, $h(i)$'s are the impulse response coefficients and $s(i)$'s are the step response coefficients. 'k' denotes the sampling instant; $t_k = kT_s$ (T_s = sampling interval).

$$y_k = \sum_{i=0}^{k} \frac{h(i)}{\Delta} \Delta u(k-i)$$

$$= \sum_{i=0}^{k} s(i)\, \Delta u(k-i)\,,$$

where

$$s(i) = \frac{h(i)}{\Delta}$$

$$= \frac{h(i)}{1 - z^{-1}} . \qquad (4.4)$$

From equation (4.4),

$$h(i) = s(i)(1 - z^{-1})$$

$$= s(i) - s(i-1) .$$

Thus, the step response coefficients can be calculated from impulse response coefficients as

$$s(i) = h(i) + s(i-1) ,$$

where

$i = 1, 2, \ldots, N$, represents the sampling instants.

Since the regression matrix X was constructed with fast sampling input data, the resulting impulse and step response models are automatically the fast-rate model with the same sampling interval of fast sampled data.

4.3 Determination of Optimal Window Size and Optimal Lag Combination

Since the FIR models are nonparsimonious, concepts of optimal window size and optimal lag combination are introduced to make the model as compact as possible. The optimal window size and optimal lag combination are determined using the least mean square error (MSE). The MSE is one of the measures of model adequacy and is a widely accepted metric in determining the fit / validity of an identified model. In DSAR, firstly the maximum window size for each of the inputs (for MISO or MIMO

systems) is determined based on the physical system and *a priori* process knowledge; the maximum window size is the past memory of the process inputs that has effect on the current value of the output. Generally, having more variables in the X matrix can give a better model fit. However, there would be room for optimization based on the concept of 'optimal window size'. In determining the optimal window size, we try to find out all possible lags from 1 to the predefined maximum window size, and then the number of lags which give the least MSE is chosen as the optimal window size. After that, the optimal lag combination is determined based on the optimal window size by fitting models using all possible lag combinations (from zero to previously determined optimal window size) for each of the inputs. The optimal combination of various lags which gives the least MSE is chosen. After performing this two stage "optimization", the unnecessary input variables and unnecessary lags of each of inputs would have been discarded.

4.4 Simulated SISO example

The continuous-time SISO system $\frac{e^{-5s}}{156.25s^2 + 18.75s + 1}$ (second order with time-delay system) is considered. The "process" was perturbed with a random input signal to generate the noise free output. To this noise free output, the output from the noise model $N = \frac{1}{0.5s + 1}$ driven by a Gaussian input sequence (variance = 0.001) was added to generate "noisy" process data. The fast-rate data set for single-rate system with 15000 observations for each input and output variable was generated with sampling interval of one time unit. The multirate systems for different integer γ values of 5, 10, 15, 20, 25 and 30 were constructed from generated single-rate system (fast-

rate data). Thus, the SISO multirate systems with one time unit sampling interval of input and slow-rate output data with different γ values were obtained. The fast-rate step response models for different γ values of multirate system were estimated using DSAR. The estimated fast-rate step response model for each γ value was compared with actual discrete-time model with one time unit sampling interval in the Figures 4.1 through 4.6. From these figures, it can be concluded that DSAR method is versatile enough for different integer γ values. Therefore, the use of DSAR for large γ values (which are very common in the chemical industries) appears to be very promising whereas the larger γ values might cause problems for the lifting method.

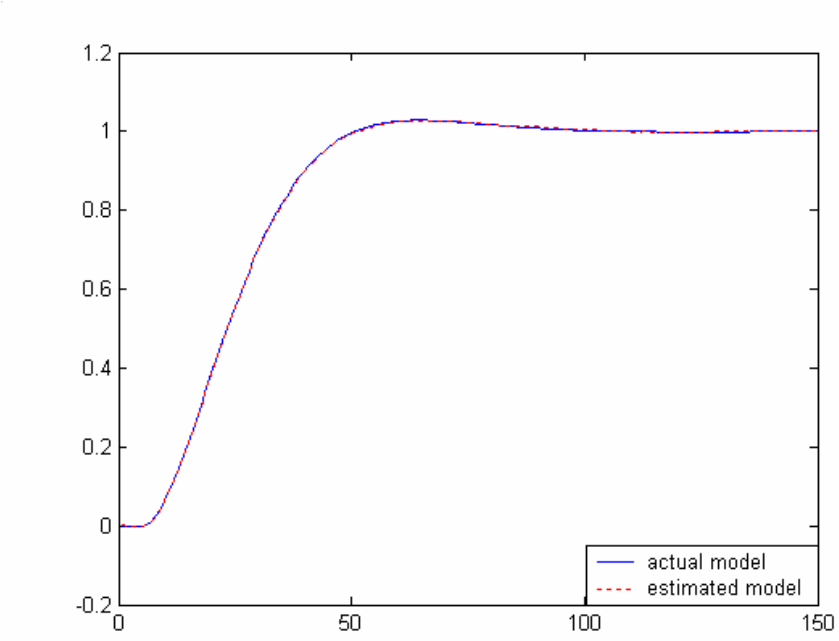

Figure 4.1: Model comparison for γ = 5

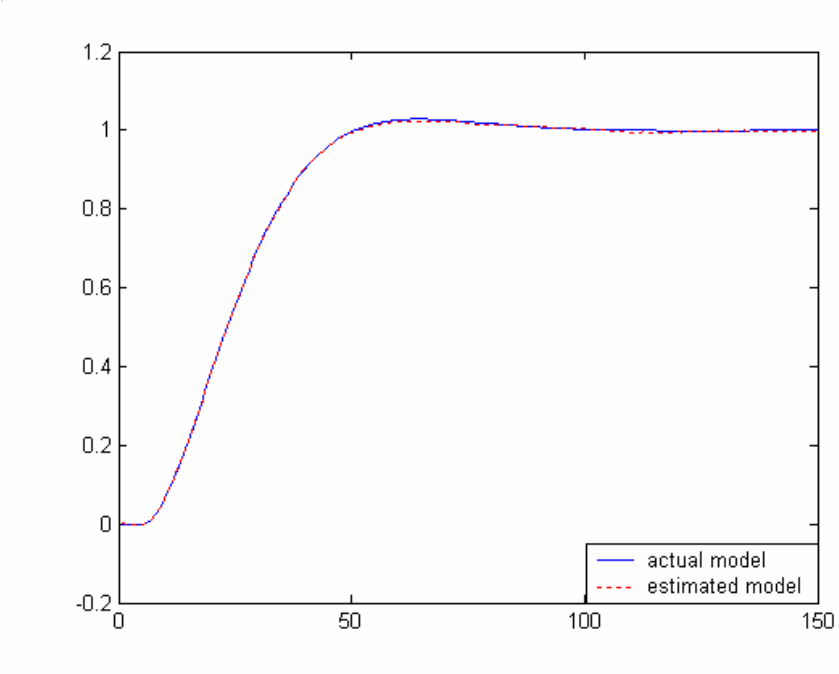

Figure 4.2: Model comparison for $\gamma = 10$

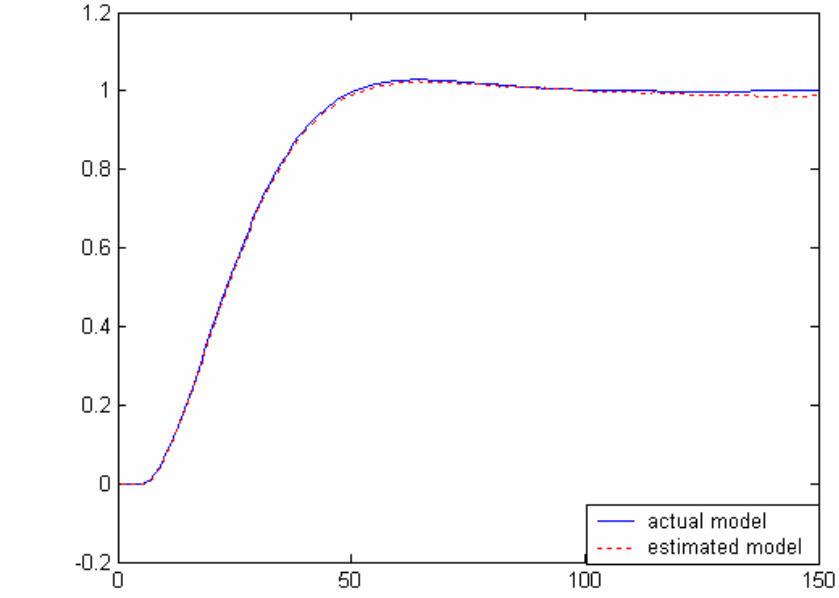

Figure 4.3: Model comparison for $\gamma = 15$

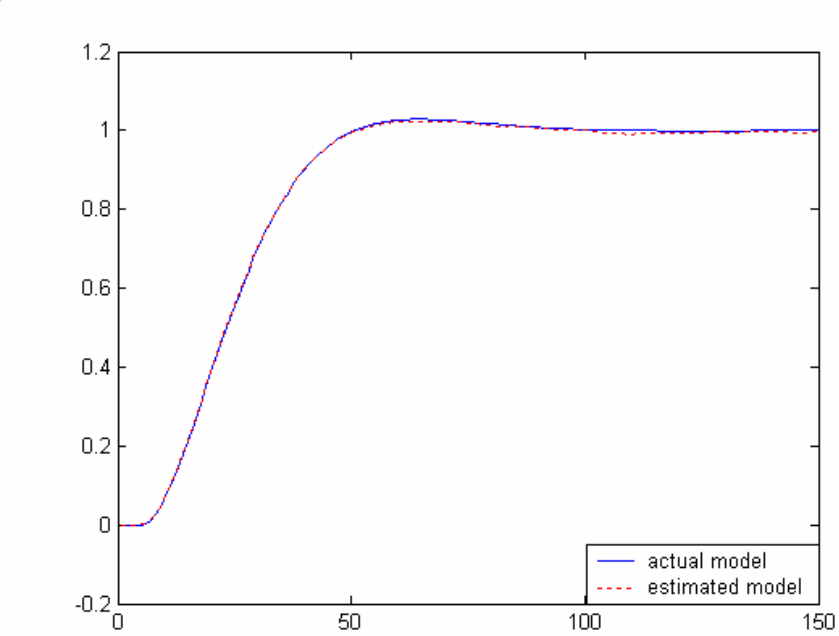

Figure 4.4: Model comparison for $\gamma = 20$

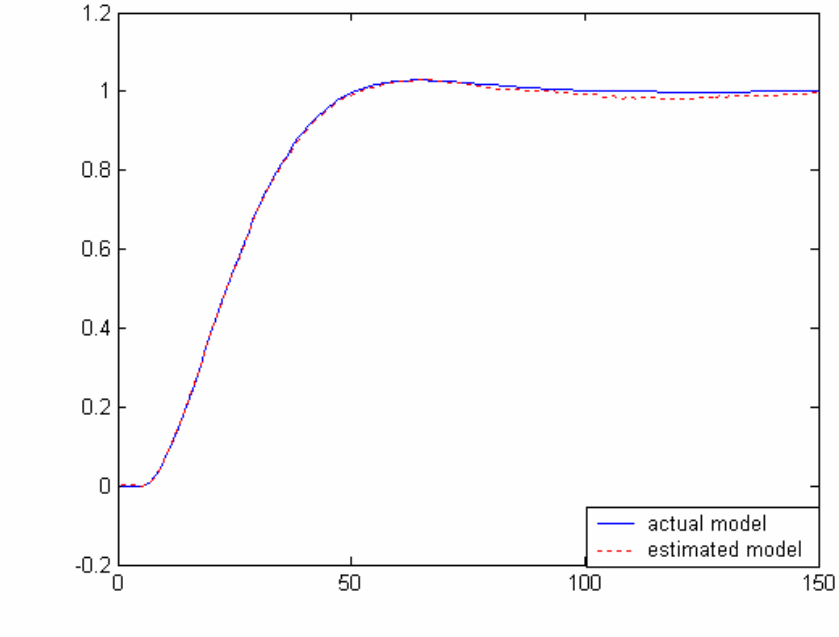

Figure 4.5: Model comparison for $\gamma = 25$

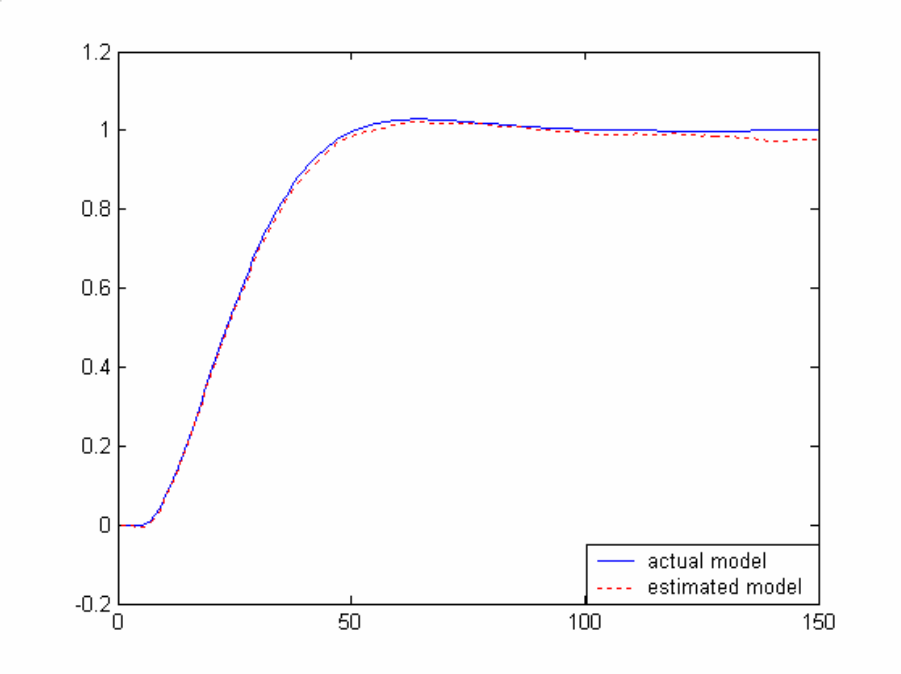

Figure 4.6: Model comparison for $\gamma = 30$

4.5 Comparison of DSAR and Lifting on University of Alberta's Data Set

The details of the data obtained from University of Alberta's experimental stirred tank system have already been discussed in Chapter 2 (section 2.5) and Chapter 3 (section 3.4). Here, the performance of DSAR and the lifting technique are evaluated on this data set. The methods of extracting the fast-rate model from identified slow-rate model were presented in section 3.3 and application of these methods to the multirate system with non-integer ratio of sampling interval was straight forward in lifting technique. Therefore, only the extraction of fast-rate model by DSAR method for the case of non-integer γ value is explored in this section.

4.5.1 Extracting of Fast-rate Model Using DSAR for Non-integer γ

It is straight forward using DSAR for the case of multirate system in which the ratio of sampling interval of input to output (γ) is integer. For the non-integer ratio of sampling intervals, it is not hard to obtain a fast-rate model with the sampling interval mp for the case of multirate system with mp sampling interval (fast control rate) and np sampling interval (slow output sampling rate) in which $m < n$; the X matrix was constructed with input data sampled at every mp time units. The first order or second order plus time-delay transfer function model is estimated using the step response coefficients which are available for mp sampling interval. Then the fast-rate model with p sampling interval can be obtained using the 'd2d' command (available in Matlab software package). The application of this "extended" DSAR method is demonstrated with the UofA data set here. Figure 4.7 shows four step response trajectories. The dash-dot and continuous plots represent the result of DSAR and CVA, respectively, using the fast data. These are almost overlapping and these serve as the standard against with the results of the multirate identification methods would be compared. The estimated fast-rate step response model using slow rate model estimated by DSAR (dotted line) and the model obtained using the lifting technique (dashed line) are also shown in the same figure. From this figure, it is observed that the performances of both of DSAR and lifting methods are quite similar for the single rate system identification. The gain of fast-rate model using DSAR is underestimated and that of the lifting technique is overestimated for this data set. Both DSAR and lifting method can capture the time constant well. In this case study, the matrix roots approach is applied for extracting the fast-rate model from the identified slow-rate model. The cross validation of both methods is also shown in Figure 4.8 and 4.9 respectively.

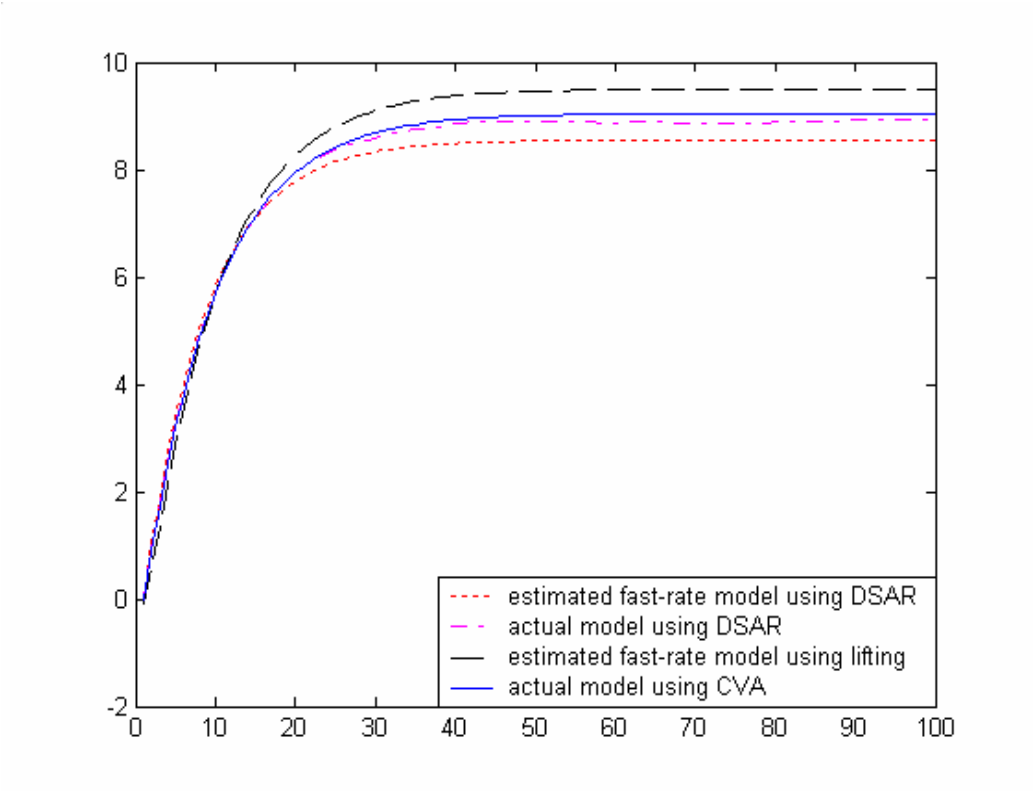

Figure 4.7: Comparison of fast-rate step response models obtained from DSAR and lifting technique

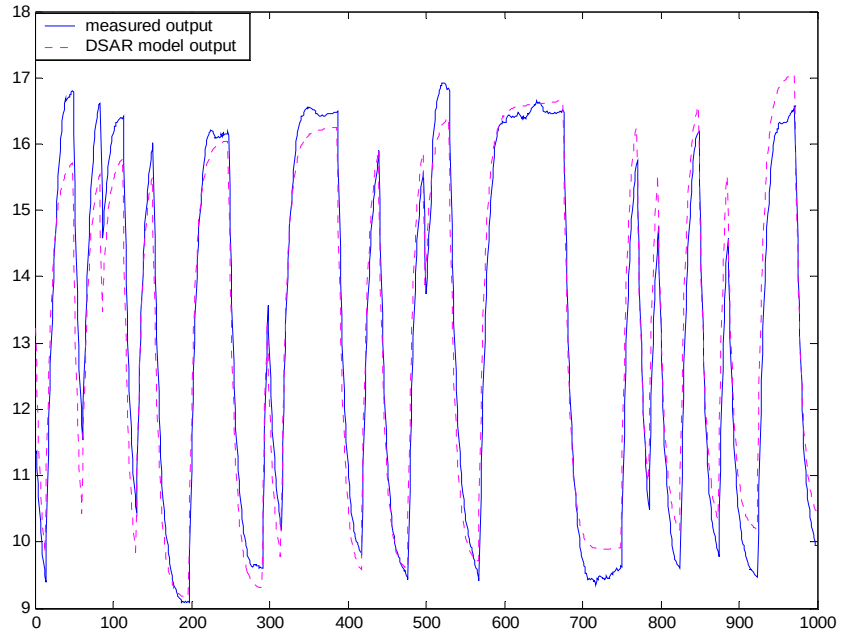

Figure 4.8: Cross validation of DSAR method

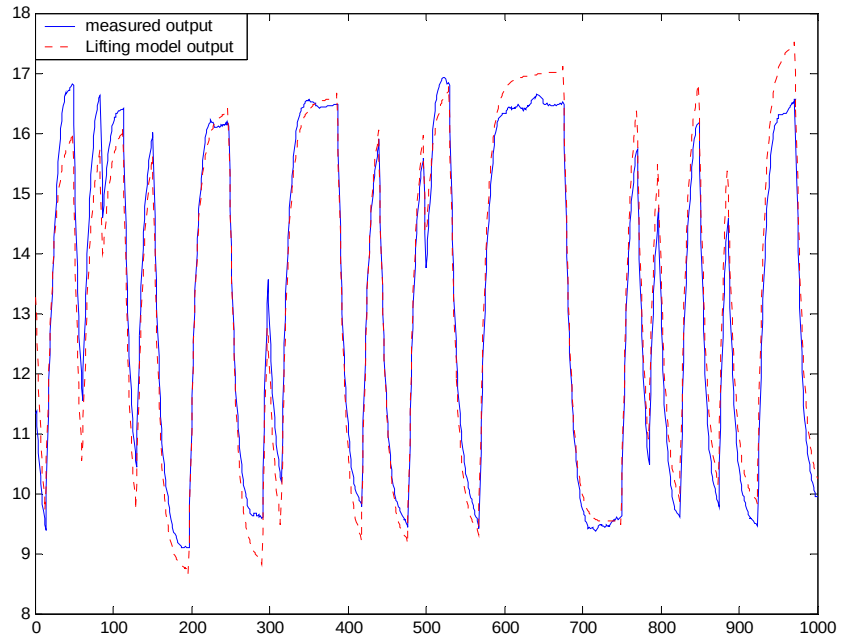

Figure 4.9: Cross validation of Lifting technique

Figures 4.8 and 4.9 confirm that both DSAR and lifting have a small mismatch in estimating the steady state gains but the process dynamics is very well captured.

4.6 Conclusions

In this chapter, a practically more useful method named DSAR for the identification of multirate system (in which the ratio of input to output sampling interval is large) is presented. This method is based on the nonparsimonious FIR model identification. The extraction of fast-rate model for integer γ is evaluated with a case study (different large γ values are provided). Moreover, the "extended" DSAR method for the non-integer γ value is also explored and this method is evaluated with application to an experimental data set (with non-integer γ value).

CHAPTER 5

CASE STUDIES OF MULTIRATE IDENTIFICATION

In this chapter, the application of the methods developed in the earlier chapters to different simulated, experimental and industrial data sets are investigated. Effects of the input signal, the sampling interval ratio (γ) and the method of identification on the quality of the identified model are studied. The effect of the value of γ on the identified model is a key characteristic that must be understood so that the method may be applied on industrial data – this will answer the question "How tolerant are these methods to the non-availability of information?" For the purpose of good multirate identification, it is important to know the best input signal to perturb the system with. These studies are done to test the measure of usefulness and to understand the limitations of the different methods for both linear and nonlinear multirate system identification.

5.1 Effect of Gamma on Linear System Identification Using Lifting Technique

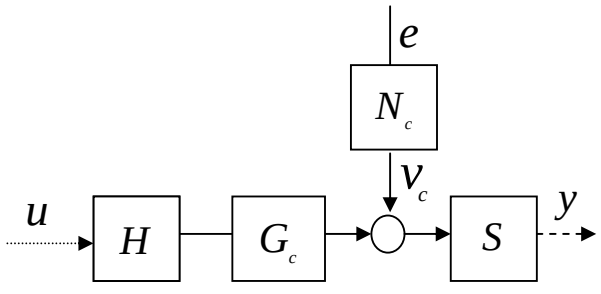

Figure 5.1: A SISO Multirate System

Figure 5.1 represents the SISO multirate sampled-data system that is used in simulation studies. Here G_c is a continuous-time LTI and causal system with or without a time delay; H is a zero-order hold with an updating period mp and S is a sampler with period np, where m and n are different positive integers, and p is a positive real number called the base period; discrete time signals u and y are the system input and output respectively; N_c is the continuous-time transfer function to which the Gaussian sequence e is introduced to produce the colored noise ν_c, which is added to the measured output to create noisy output data. In particular, we consider a system where $G_c(s) = \frac{2.4}{112.5s^2 + 21.5s + 1}$ and $N_c(s) = \frac{1}{0.5s + 1}$. The Gaussian signal with variance = 1 is designed as input signal (u) and e is a white noise sequence with variance = 0.001. 9000 input/output data pairs are generated at every one time unit – these data can provide us the reference fast-rate model for comparison purposes. To study the effect of γ on the identification of fast-rate models which are extracted from different multirate data sets with different γ values, input data are collected at every one time unit and output data are collected at every γ time units sampling interval. Thus, $m = 1, p = 1, n = \gamma$ are used for simulation in this case. Multirate data with different γ values (γ = 4, 8, 12, 16) were created. The fast rate model (for one unit sampling interval) is extracted from the estimated slow rate models and the estimated step response of the fast rate models are compared with the actual step response. In Figure 5.2, the result of the lifting technique for different γ values is provided. As expected, the quality of the identified model becomes poorer with increasing γ value. In particular, it is found that time constant can be estimated properly but the model gain cannot be estimated exactly – this trend is evident with higher γ values.

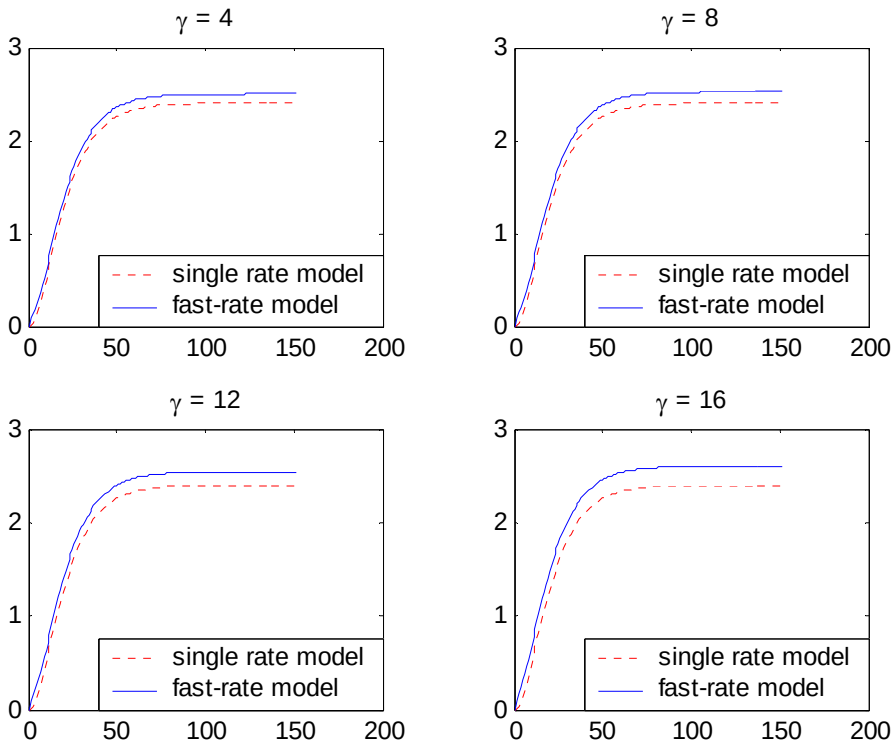

Figure 5.2: Comparison of single-rate and fast-rate model using lifting technique

Other case studies were performed with different input signals like as PseudoRandom Binary Sequence (PRBS), stretched PRBS and random signal with uniform distribution. From these case studies, it is found that lifting technique is versatile enough for different kind of input signals for linear system identification. The trends shown above persisted even with these input signals.

5.2 Effect of Gamma on Nonlinear System Identification using Lifting Technique

5.2.1 Hammerstein Model Multirate System Identification

5.2.1.1 SISO Hammerstein Model MRID

For SISO Hammerstein model multirate system identification, the simulated SISO system in which the linear dynamic part follows the static nonlinearity was built. For

the static nonlinear part, the polynomial $0.2u^3 + 0.3u^2 + u$ was assigned. The first order discrete-time model $\frac{0.2}{z - 0.8}$ was assigned as the linear dynamic part in the simulated system. The random signal with normal distribution was designed as input signal. The 2000 input/output single-rate data were collected with one time unit sampling interval as the reference fast-rate data set. White noise was added to the noise free output signal to simulate an output sequence with a signal to noise ratio of 10. The different multirate data sets were collected with one time unit of input sampling interval and (one*γ) time unit of output sampling interval; γ values used for this study are 1, 2, 3, 4, and 5. The fast-rate model was extracted from each constructed multirate systems using the developed Hammerstein model multirate system identification algorithm (section 3.5.2) with the matrix roots approach. For the identification of slow-rate linear dynamic subsystem, MOESP scheme developed for single-rate Hammerstein model identification (Verhagen & Westwick (1996)) was used (SLS was used for the identification of whole system). The cross validation of the estimated fast-rate model output with measured fast-sampled output was performed. The cross validation of the models for the different γ values are shown in Figures 5.3 to 5.7. From these figures, it can be seen that the estimated fast-rate model is quite acceptable for different γ values for the multirate system with random input signal. Persistency of excitation provided by the random input signal and the relatively good signal to noise ratio are perhaps why good identification results are obtained. Therefore, the model outputs from the estimated fast-rate models of the different γ values are quite identical to the actual fast-sampled data. From this study, it seems that the γ value has little impact on the extraction of fast-rate model if the system is perturbed with a persistently exciting signal. It appears that the random

signal may be the best perturbation signal to achieve a good model even in the multirate system identification scenario.

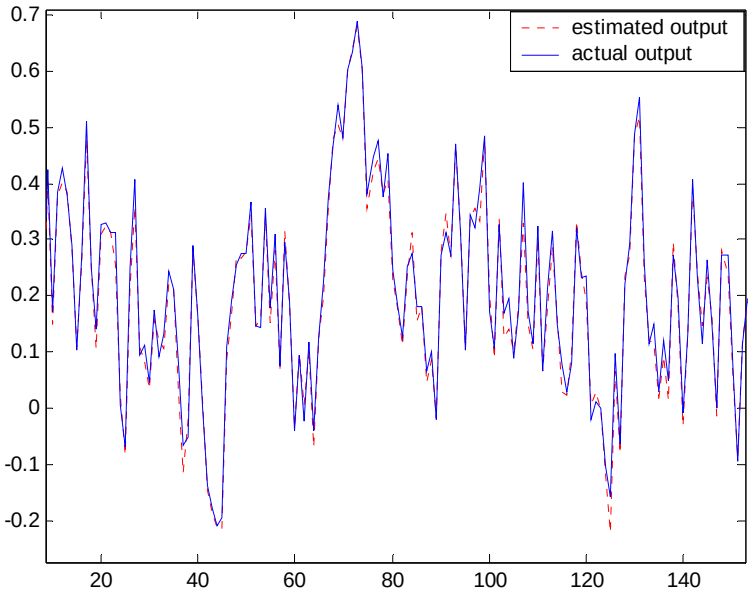

Figure 5.3: Cross validation for $\gamma = 1$, H- type SISO MR System

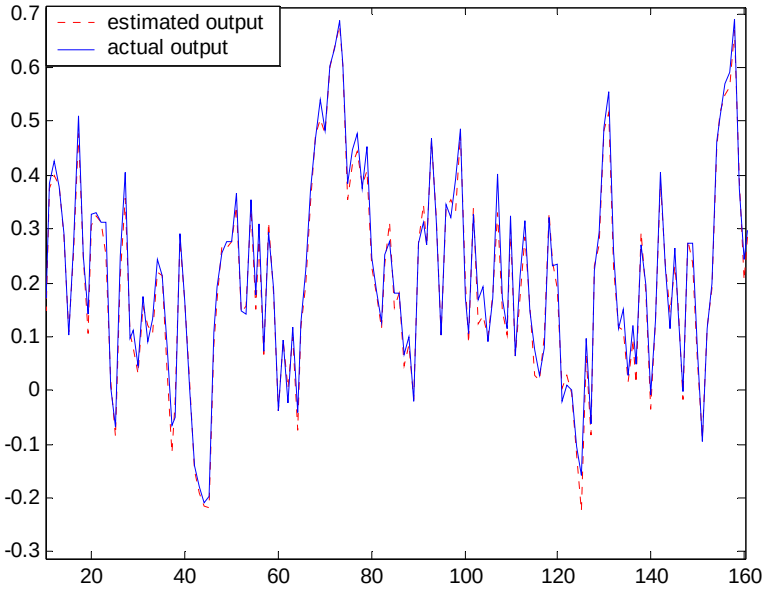

Figure 5.4: Cross validation for $\gamma = 2$, H- type SISO MR System

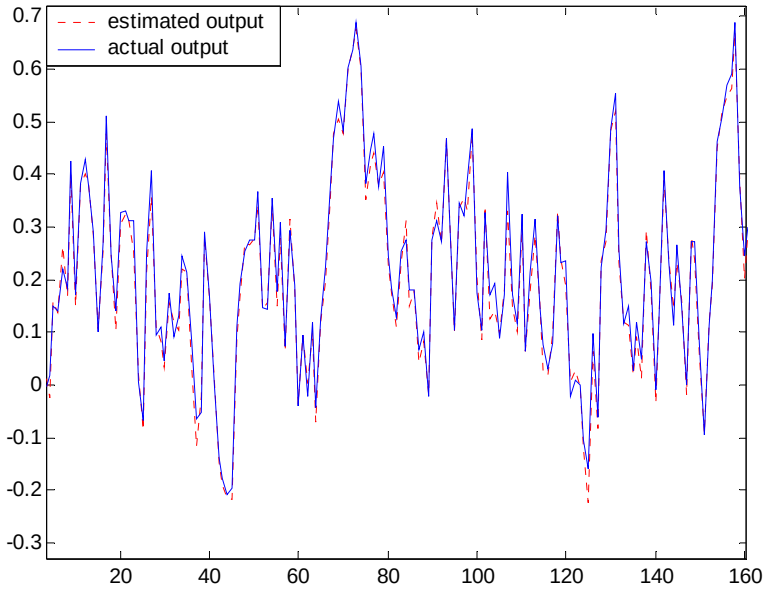

Figure 5.5: Cross validation for $\gamma = 3$, H- type SISO MR System

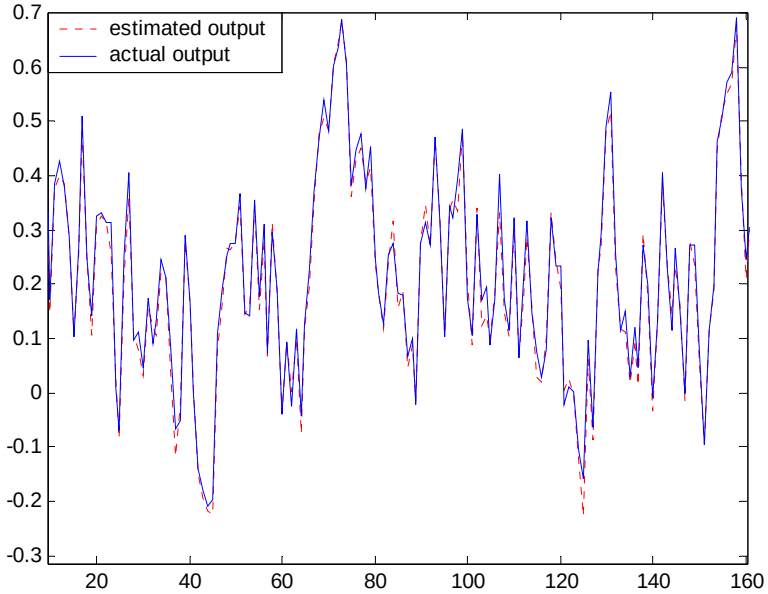

Figure 5.6: Cross validation for $\gamma = 4$, H- type SISO MR System

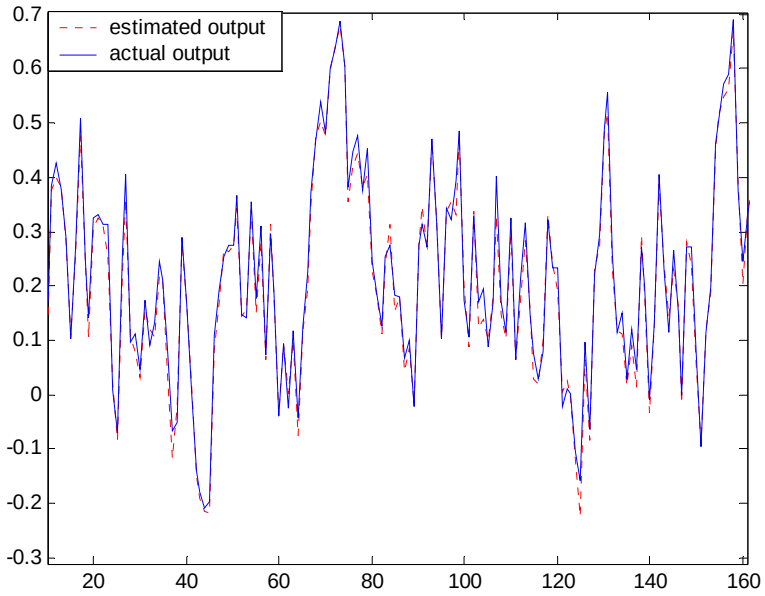

Figure 5.7: Cross validation for γ = 5, H- type SISO MR System

5.2.1.2 MISO Hammerstein Model MRID

The MISO discrete-time Hammerstein model was built by using the two different inputs each have different model dynamic and different static gain and then these two inputs combined as one system output; the output was produced by the first input channel which perturb the static nonlinearity $0.3u_1^2 + 0.6u_1$ followed by dynamic subsystem $\frac{0.1538}{z-0.9231}$, and the second input channel which pass through the static nonlinear subsystem $0.3u_2^3 - u_2^2 + 0.3u_2$ and linear dynamic part $\frac{0.2}{z-0.8}$. The system inputs were random signals with normal distribution and which were produced at different states, so that they are different sequences. Noise was added to the system outputs so as to achieve a noise to signal ratio of 0.1. This is done in order to mimic

the characteristics of real world data sets. The fast sampled single-rate data set (in which 2000 input/output data pairs) was collected with one time unit sampling interval. This fast-rate data set was used for cross validation purpose (to measure the adequacy of extracted fast-rate model output to the actual fast sampled system output). The different multirate data sets were collected from this fast sampled single-rate data set, so that the constructed multirate systems have fast sampled inputs which are sampled at every one time unit and the slow sampled output which has (one*γ) sampling interval. These kind of multirate systems were built for different γ values from 1 to 5. The γ value 1 was used as the special case of multirate system. Then the fast-rate models were extracted from the different multirate data set by employing the same algorithm (Hammerstein model multirate system identification method). The adequacy of estimated fast-rate model for each γ value was measured by comparing the estimated fast-rate model output with the actual fast sampled system output. As with SISO Hammerstein model multirate identification, the effect of γ on the identification results is not significant at least with random probing signals. From these observations, it would appear that random input signal could be the best for the identification of Hammerstein models from multirate data.

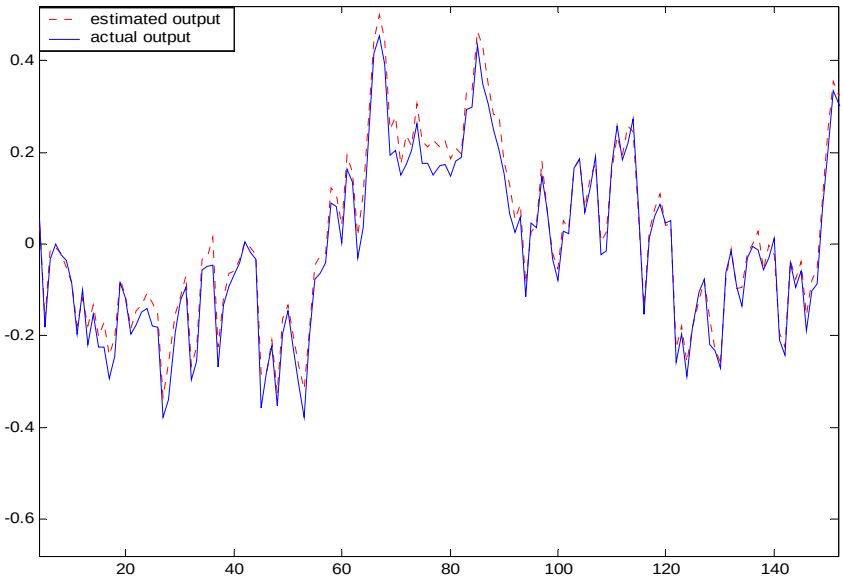

Figure 5.8: Cross validation for $\gamma = 1$, H- type MISO MR System

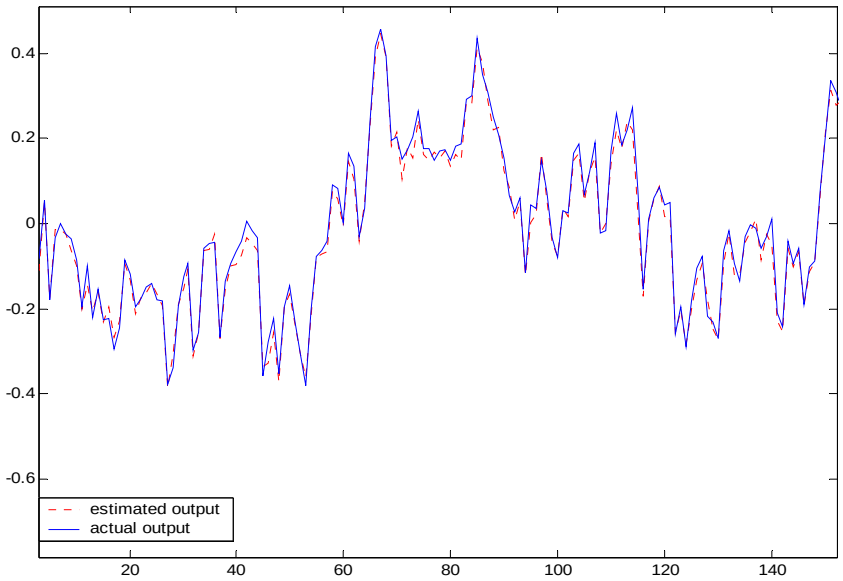

Figure 5.9: Cross validation for $\gamma = 2$, H- type MISO MR System

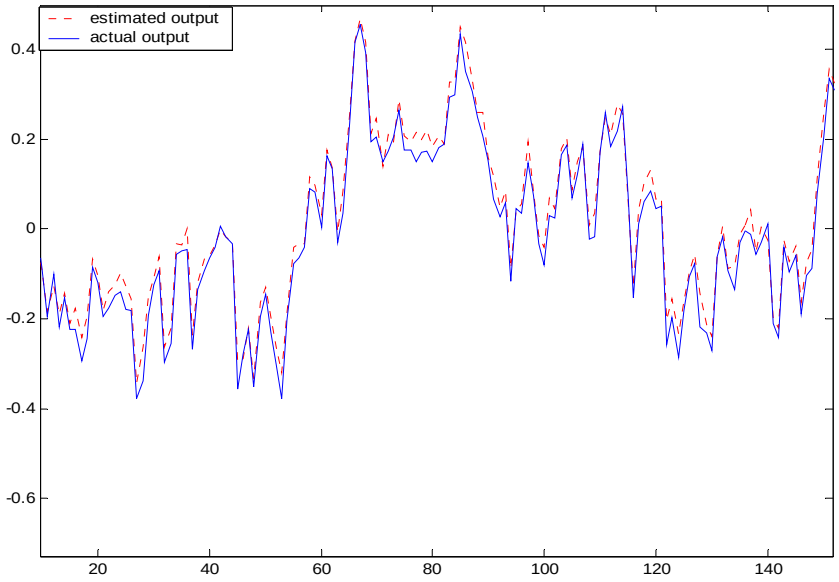

Figure 5.10: Cross validation for $\gamma = 3$, H- type MISO MR System

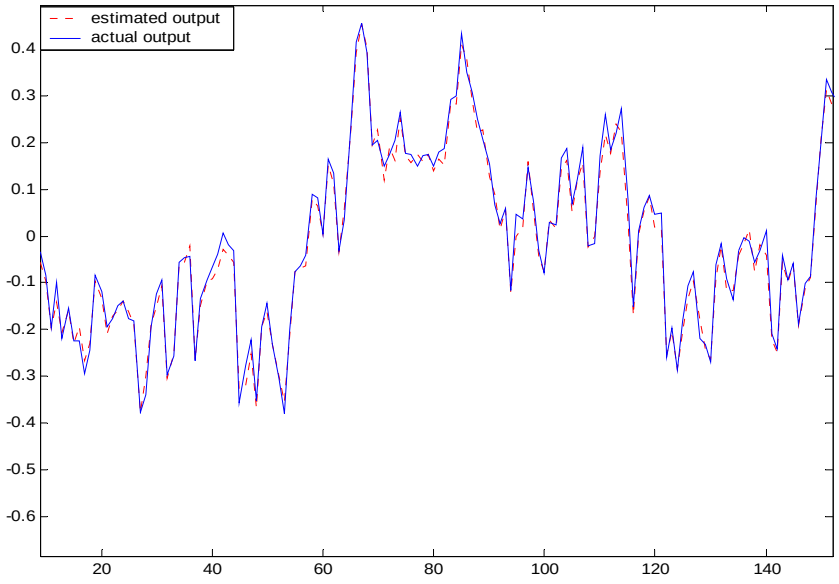

Figure 5.11: Cross validation for $\gamma = 4$, H- type MISO MR System

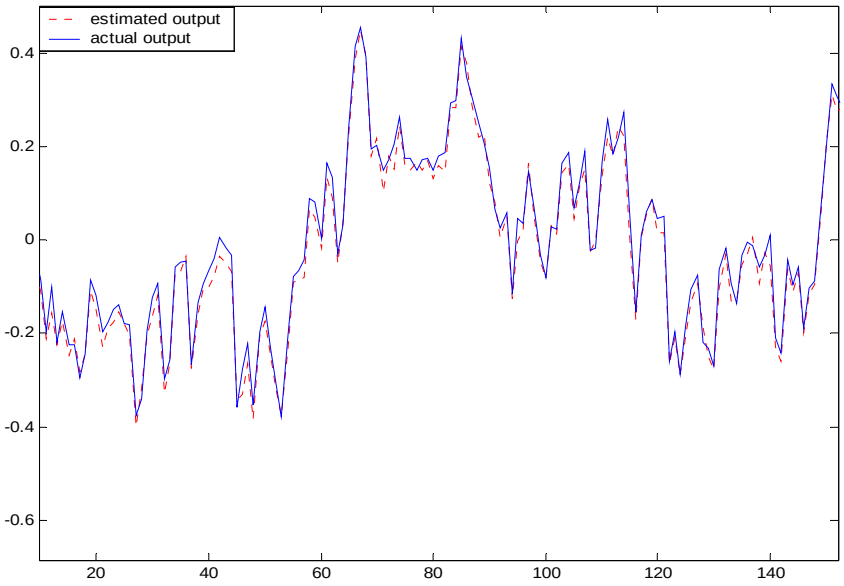

Figure 5.12: Cross validation for γ = 5, H-type MISO MR System

5.2.2 Wiener Model Multirate System Identification

5.2.2.1 SISO Wiener Model MRID

Since the Wiener Model is the reverse of Hammerstein Model, the simulated SISO system in which the linear dynamic subsystem is followed by the static nonlinearity was built for generating identification data. The first order discrete-time LTI system $\frac{0.2}{z-0.8}$ was assigned as the linear dynamic part in the simulated system. For the static nonlinear part, the polynomial $0.6u + 0.3u^2 - 0.2u^3$ was assigned. The random signal with normal distribution was chosen as the input signal to this simulated nonlinear process. 2000 input/output single-rate data were collected with one time unit sampling interval as the reference fast-rate data set. Noise was added to the output data to create an output sequence that had a signal to noise ratio equal to 10. The different multirate

data sets were collected with one time unit of input sampling interval and (one*γ) time unit of output sampling interval; γ values used for this study are 1, 2, 3, 4, and 5. The fast-rate model was extracted for each case (different γ) using the developed Wiener model multirate system identification algorithm (section 3.5.3) with the matrix roots approach. For identification of the slow-rate linear dynamic subsystem, MOESP scheme developed for single-rate Wiener model identification (Verhaegen & Westwick,1996) was used (SLS was used for the identification of whole system); the initial estimates are calculated by using the MOESP scheme ((Verhaegen & Westwick,1996) and then they are subsequently improved by using iterative optimization (Bruls et al. (1999)). The identified nonlinear static subsystem is in Tchebyshev polynomial form. The cross validation of the estimated fast-rate model output with measured fast-sampled output was performed. Figure 5.13 to Figure 5.17 depict the cross validation results for models obtained with different γ values. From these figures, it can be seen that the estimated fast-rate model is quite acceptable for different γ values for the Wiener type multirate system with random input signal. The model outputs from the estimated fast-rate models with the different γ values are quite identical to the actual fast-sampled data. This shows the utility of the identification strategy proposed in this work.

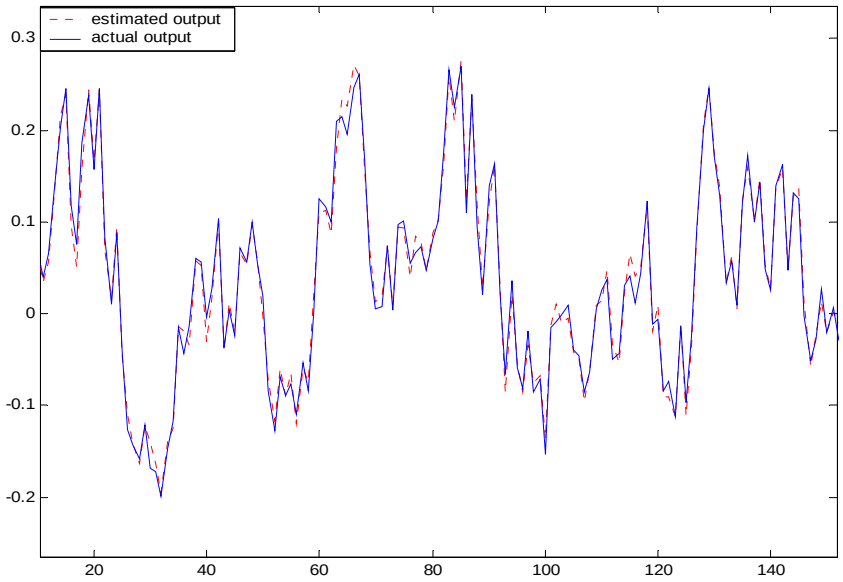

Figure 5.13: Cross validation for $\gamma = 1$, W-type SISO MR system

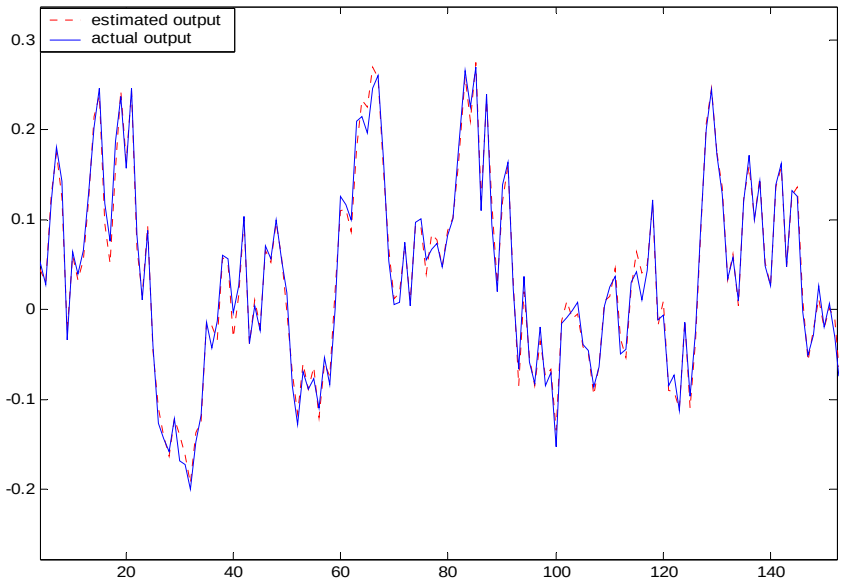

Figure 5.14: Cross validation for $\gamma = 2$, W-type SISO MR system

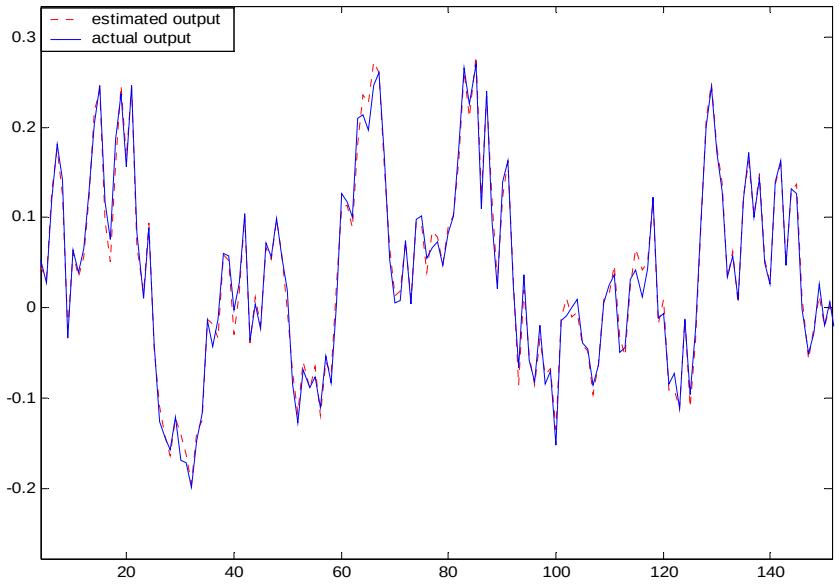

Figure 5.15: Cross validation for $\gamma = 3$, W-type SISO MR system

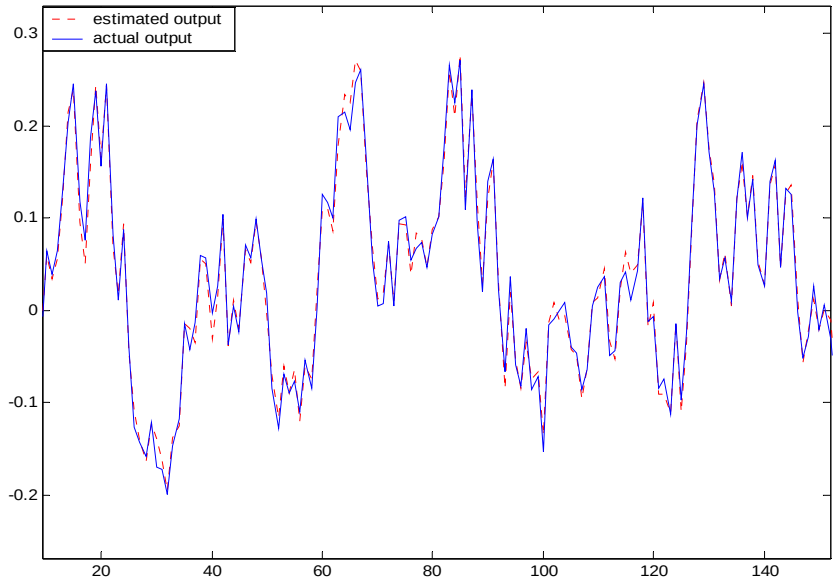

Figure 5.16: Cross validation for $\gamma = 4$, W-type SISO MR system

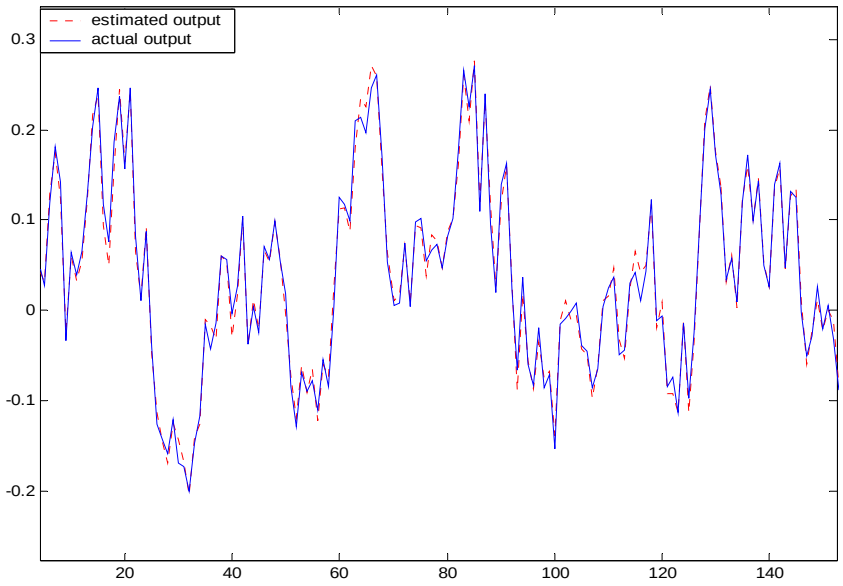

Figure 5.17: Cross validation for γ = 5, W-type SISO MR system

5.2.2.2 MISO Wiener Model MRID

The MISO discrete-time Wiener model was built by using the two different inputs each have different model dynamic and different static gain, and then these two inputs combined to form only one system output; the output was produced by the two input channels which have the same static nonlinearity $0.25u + 0.25u^2$ (which is the Tchebyshev kind of polynomials) followed by different dynamic subsystems, $\frac{0.1538}{z - 0.9231}$ (for first input channel) and $\frac{0.2}{z - 0.8}$ (for second input channel). The system inputs were random signals with normal distribution and which were produced at different states, so that they are different sequences. Random noise was added to the noise-free output data to generate a measured output sequence with a noise to signal

ratio is 0.1. The fast sampled single-rate data set (2000 input/output data pairs) was collected with one time unit sampling interval and which was used as the data set for cross validation purposes (to measure the adequacy of extracted fast-rate model output to the actual fast sampled system output). The different multirate data sets were collected from this fast sampled single-rate data set, so that the constructed multirate systems have fast sampled inputs which were sampled at every one time unit and the slow sampled output which were sampled at (one*γ) sampling interval. These kind of multirate systems were built for different γ values from 1 to 5. The γ value 1 was used as the special case of multirate system. Then the fast-rate models were extracted from the different multirate data set by employing the same algorithm (Wiener model multirate system identification method). The adequacy of estimated fast-rate model for each γ value was measured by comparing the estimated fast-rate model output with the actual fast sampled system output as shown in Figures 5.18 to 5.22. As in the SISO Wiener model multirate system, the effect of γ is not significant for the system with random probing signals. From these figures, the cross validation model fits are quite acceptable but are probably not so good as in SISO case (see Table 5.1). From these observations, random signal could be the best for the Wiener model multirate system identification as with Hammerstein model identification.

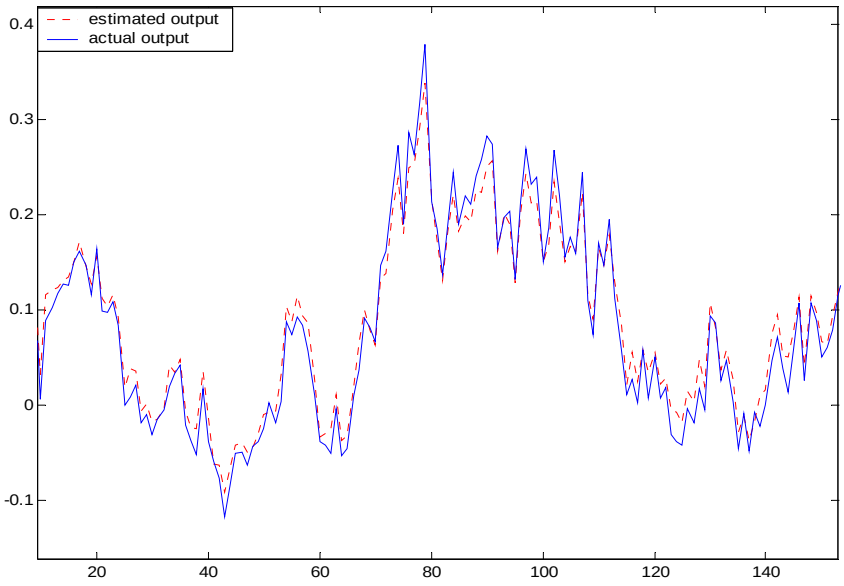

Figure 5.18: Cross validation for $\gamma = 1$, W-type MISO MR system

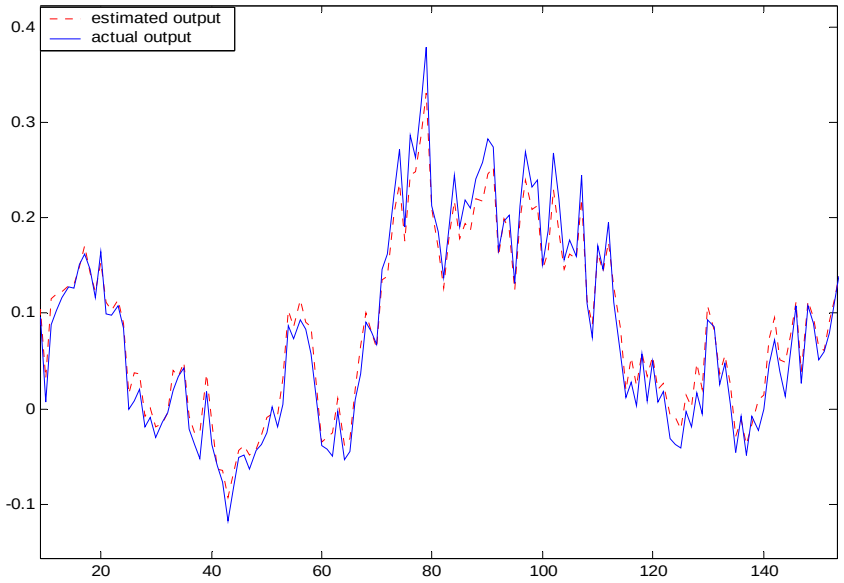

Figure 5.19: Cross validation for $\gamma = 2$, W-type MISO MR system

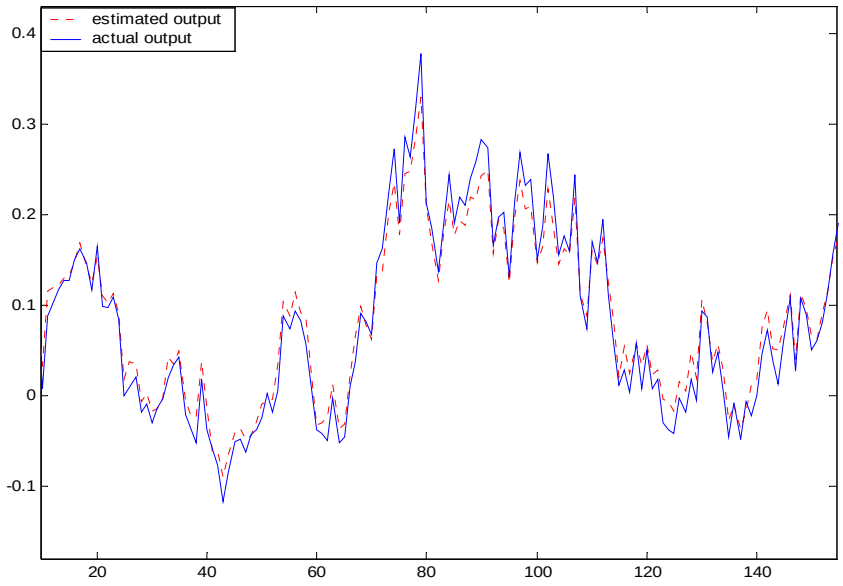

Figure 5.20: Cross validation for $\gamma = 3$, W-type MISO MR system

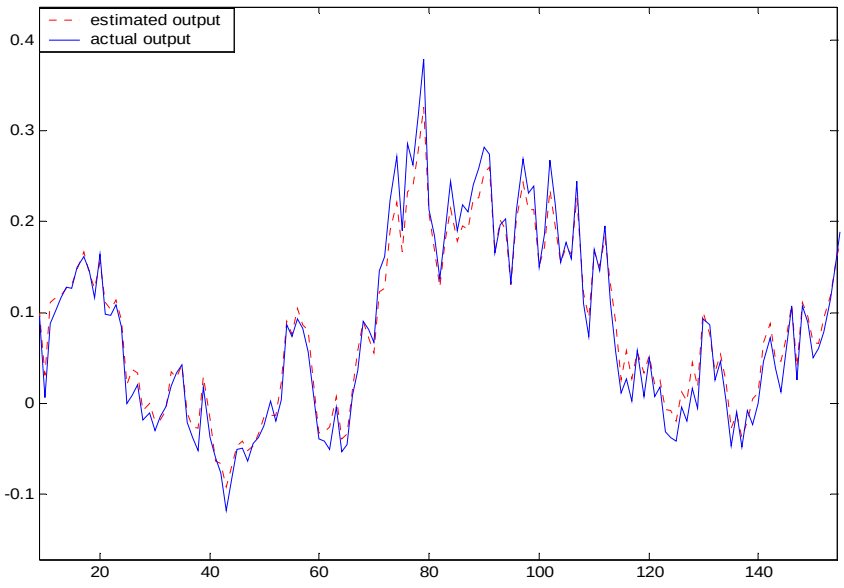

Figure 5.21: Cross validation for $\gamma = 4$, W-type MISO MR system

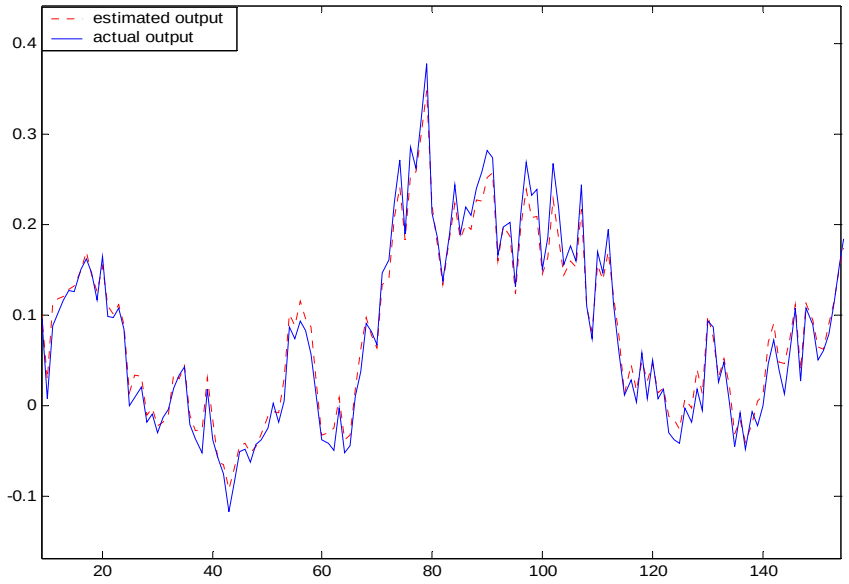

Figure 5.22: Cross validation for γ = 5, W-type MISO MR system

5.2.3 Effect of Gamma in MSE criteria

Table 5.1. Mean square error values for both SISO & MISO multirate system of H-type and W-type model

γ	H-type SISO	H-type MISO	W-type SISO	W-type MISO
1	1.1166	2.3672	0.2643	0.5756
2	1.1960	1.1649	0.2654	0.5702
3	1.1796	2.1114	0.2651	0.6084
4	1.2239	0.9064	0.2657	0.5929
5	0.9713	1.9551	0.2724	0.4158

The mse values for different γ values are summarized in Table 5.1 for the case studies performed in sections 5.2.1.1, 5.2.1.2, 5.2.2.1, and 5.2.2.2. These mse values are

calculated based on the difference between estimated fast-rate model fitness and the actual fast sampled output. From this table, it can be seen that mse does not increase with increasing gamma value. It may be due to convergence of the optimization problem to local optima (since nonlinear optimization algorithm and solver used here are only capable of finding local optima (according to Edgar et al. (2001)).

5.3 Effect of Input Signals on DSAR Identification

It was demonstrated in section 4.4 that the DSAR method is robust for various γ values. However, the effect of different kind of input signals on the DSAR technique needs to be quantified. This is the objective pursued here.

As shown in Figure 5.1, the SISO simulation model was built using Simulink toolbox in Matlab; the second order continuous-time linear time invariant (LTI) model $G_c = \frac{2.4}{156.25s^2 + 21.75s + 1}$ was used as the process and $N_c = \frac{1}{0.5s + 1}$ was used as the disturbance model. A maximum length PRBS signal was generated and stretched by a stretch factor of 17. The stretch factor was calculated from *a priori* knowledge of the process. The fast sampled single-rate input/output data set was generated with one time unit sampling interval and 2000 input/output data points are available. The multirate systems for different γ values starting from 1 to 12 were constructed from fast sampled single-rate data set by collecting the fast sampled input signal and slow sampled output data with sampling interval equal to γ. The fast-rate model was extracted from the constructed multirate data sets by using the DSAR method. The expected time to steady state was set at 100 for the different γ values. The discrete-time step response model for one time unit sampling interval was obtained by converting the impulse response coefficients to step response coefficients.

The estimated step response model obtained from DSAR method was compared with the actual step response of the simulated process. The resulting fast-rate step response models were smooth up to $\gamma = 6$. Some estimated step response models for $\gamma = 7$, 9, 10, and 11 are not smooth, and the fast-rate step response models that are directly estimated by DSAR method for these γ values are shown in Figures 5.23, 5.24, 5.25 and 5.26 in which they are compared with actual step response of the process, respectively. For the case in which the input signal is PRBS type, the DSAR generated models must be regularized (unsmooth or noisy step responses should be smoothened using the regularization method to render them meaningful and useful). The continuous-time transfer function with time delay was estimated from unsmooth step response resulting from the DSAR method. The step response was then calculated using this continuous-time LTI system with sampling interval one time unit so that the resulting step response model would reflect the fast rate behavior of the system. These step response models for γ values 7, 9, 10 and 11 are shown in Figures 5.23, 5.24, 5.25 and 5.26 also and are compared with the step response model before performing the regularization (raw models) and true model of the system. It is found that the step response models after regularization are quite acceptable but they are not exact enough as the actual (or true) step response of the process.

From this experience, it may be surmised that if stretched PRBS signals are employed for multirate identification, some regularization of the resulting model is needed. This method is useful with PRBS kind of input signal but it must be maximum length PRBS. This method needs a well-excited input signal. DSAR is compatible for various ratios of sampling periods (both integer and non integer ratio). Moreover, from our observations with other kinds of signal, it can be concluded that DSAR method is suitable for Gaussian and random input signals also.

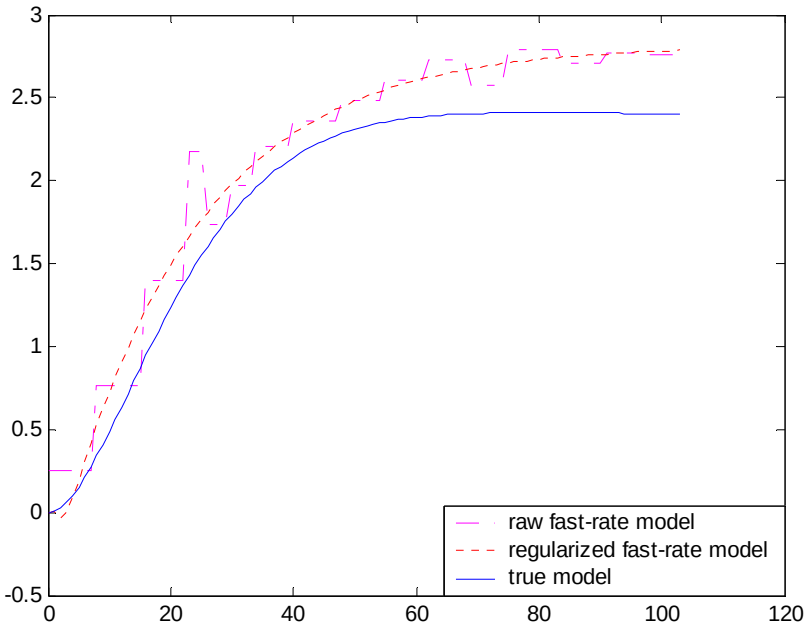

Figure 5.23: Comparison of step response models for $\gamma = 7$

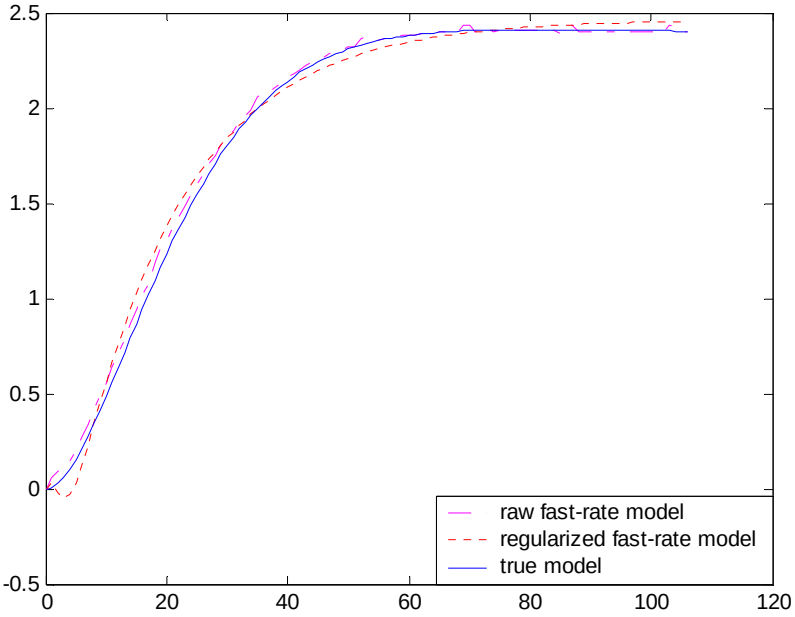

Figure 5.24: Comparison of step response models for $\gamma = 9$

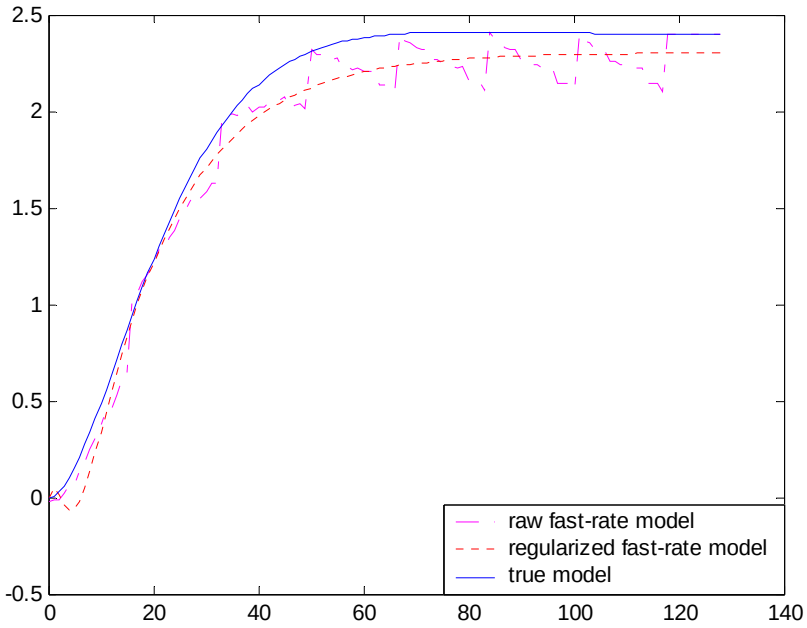

Figure 5.25: Comparison of step response models for $\gamma = 10$

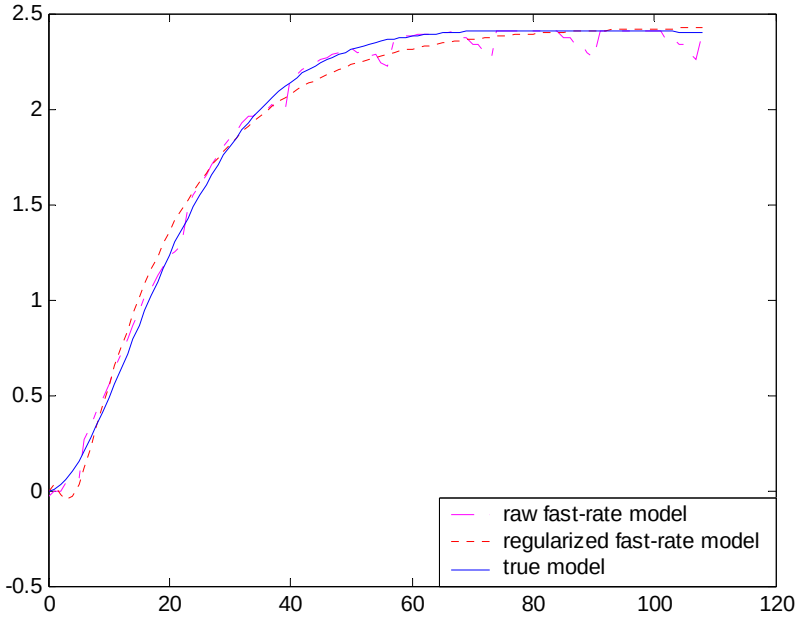

Figure 5.26: Comparison of step response models for $\gamma = 11$

5.4 DACS Experiment Data Analysis

The experimental data set from the DACS lab experimental setup was considered next. The details of this experimental setup and schematic diagram were mentioned in section 2.5.1.3 (Case Study III of Chapter 2). In this experiment, our focus is on tank 1. The experiment was conducted for the heating system, in which the heating power was considered as the input to the system and the exit water temperature was regarded as the output of the system. The flow rate was kept constant and the dynamics of the heating system was study. The input signal was designed as a multilevel signal. This fast sampled input is plotted in Figure 5.27. The single-rate input and output data were collected every one minute and this data set was applied as the reference fast sampled data set. The multirate data sets for different γ values (1 to 5) were constructed from fast sampled reference data set by collecting the fast sampled input data with one minute sampling interval and slow sampled output data with γ minutes sampling interval for each γ value. Thus, the multirate data sets for γ value 1 to 5 were obtained to perform multirate system identification. These data sets were identified with lifting technique for linear system identification in which the matrix roots approach was used to extract the fast-rate model from constructed multirate data sets. It was found that this approach failed to identify any acceptable model for the process. The modified alternate approach (section 3.5.1) was successful in identifying the fast rate model. The process was successfully identified using the Hammerstein model multirate system identification method developed in this work. The fast rate model was identified using the modified alternate approach (section 3.5.1). The process is identified as a Hammerstein model in which a first order LTI model follows the static nonlinear part that is third order polynomial. The cross validation is performed for different estimated fast-rate models; the estimated fast-rate model output is compared

with measured fast sampled output data in Figures 5.28 to 5.32 to test the model adequacy for each γ value. The mean square error (MSE) of the estimated fast-rate model output from measured fast sampled output data was also calculated to measure the model quality. It is found that the Hammerstein model is better than the linear model - this implies a nonlinear relationship between the heating power and the exit water temperature. The comparison of MSE for linear lifting technique and Hammerstein model is shown in Table 5.2.

Table 5.2. Mean square error comparison for DACS experimental data

γ	Linear lifting	Hammerstein
1	720.0	491.3
2	719.5	515.0
3	722.8	611.0
4	722.8	686.0
5	723.9	642.9

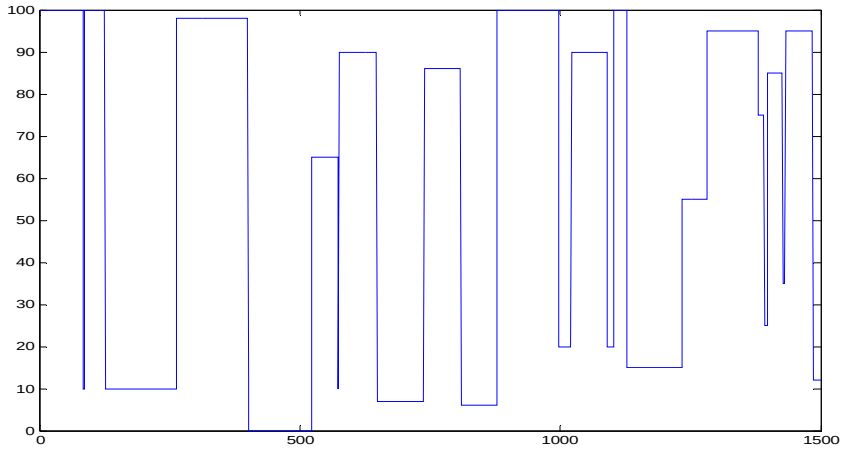

Figure 5.27: Plot of Input data for DACS data set

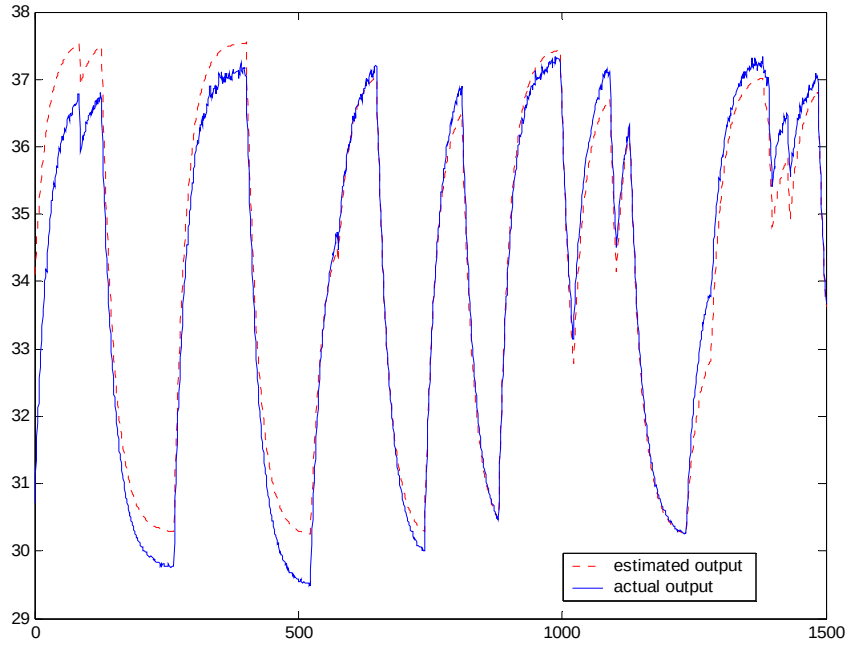

Figure 5.28: Cross validation for $\gamma = 1$, DACS data set

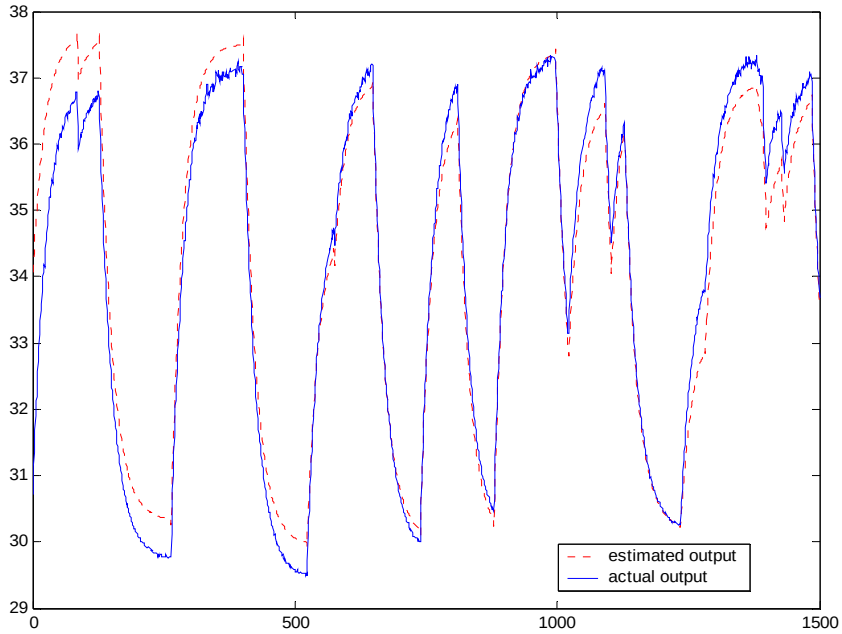

Figure 5.29: Cross validation for $\gamma = 2$, DACS data set

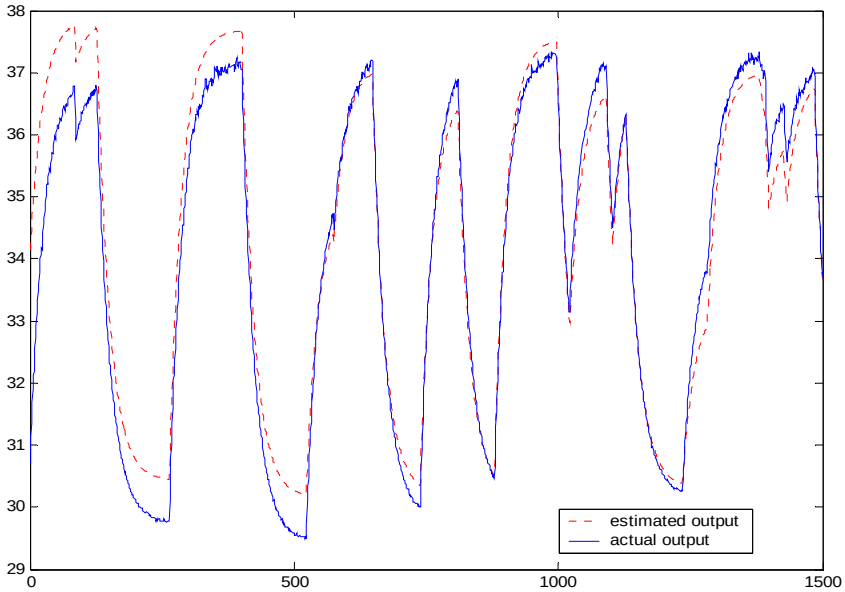

Figure 5.30: Cross validation for $\gamma = 3$, DACS data set

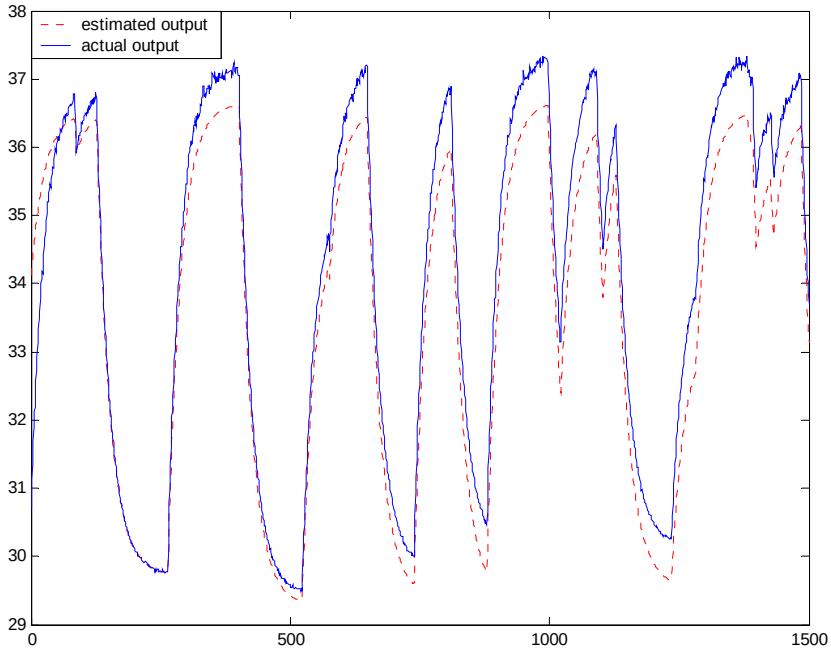

Figure 5.31: Cross validation for $\gamma = 4$, DACS data set

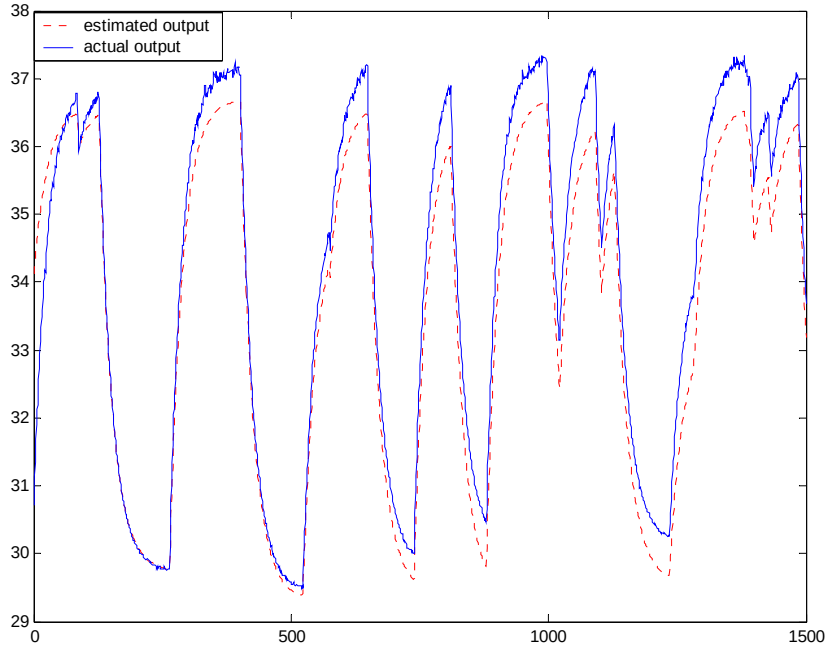

Figure 5.32: Cross validation for $\gamma = 5$, DACS data set

5.5 Industrial Application of DSAR

Since DSAR method is versatile enough for large γ values, it was chosen to identify a data based model for an industrial reactor. The data set was obtained from Mitsubishi Chemical Corporation, Mizushima, Japan. The process is the acetylene converter which is a train of two reactors whose primary function is to convert as much acetylene (from the de-ethanizer overhead) to ethylene so that the product meets stringent specification on acetylene levels. In addition, the acetylene is to be converted to ethylene and not to any non-profitable byproducts. The acetylene concentration is infrequently measured using an analyzer at the outlet of the converter (once every 40 minutes). This infrequent measurement can often lead to poor operation of the process. In this study, we examine the development of a soft sensor whose goal is to predict the outlet acetylene concentration using available information from the frequently measured variables such as flow rates, temperatures etc. The applied procedures for the development of the soft sensor are mentioned below.

5.5.1 Optimal Window Size

The Acetylene plant data (from 4/01/2003 to 7/31/2003) are divided into five data sets as follows:

SET1: from 4/01/2003 to 4/30/2003

SET2: from 5/01/2003 to 5/22/2003

SET3: from 6/01/2003 to 6/16/2003

SET4: from 6/17/2003 to 6/30/2003

SET5: from 7/01/2003 to 7/31/2003

Some outliers and spikes were omitted and each data set was differenced. Six variables were chosen from various inputs of the system. These were used as explanatory variables in the model. The past 10 samples for the first five variables (see Table 5.4) and one past sample for the sixth variable (this was a infrequently measured input variable - inlet acetylene to reactor 1; see Table 5.4) were considered as the input variables (a total of 51 "input variables"). Then the model was identified with window sizes ranging from 10 to 5, by using one data set. After that, the model was tried on the other four data sets to check out its predictive capability. The sum of mean square error (msesumw) between the predicted data and measured data (after taking difference) was evaluated. This procedure was applied in turn to every data set. The mse values for the prediction of other four data sets by a particular data set and their corresponding optimum window sizes are shown in Table 5.3. The optimum window size was chosen based on the model that provided the least msesumw value.

Table 5.3. Mean square error of various data sets

Model	msesumw	Optimum window size
SET1	1259.3	7
SET2	1307.9	7
SET3	1549.1	10
SET4	1256.3	6
SET5	1330.16	10

It is found that the model obtained from SET4 is the best in predicting the outputs of the other data sets and it has minimum window size. Thus, SET4 was selected to estimate the model and optimum window size was determined as 6.

5.5.2 Optimal Lag Combination

The model was constructed using SET4 with the lags ranging from 6 to 0 and validation on other four data sets was performed. The best optimal lag combination is determined by the model associated with least sum of mean square error (msesum) and it is shown in Table 5.4. From this analysis, it was found that we can omit two variables, u3 and u5; we can reduce the number of "input variables" to 13 (from 51) and the sum of mean square error in prediction to 1208.8 (reduction from 1256.3).

Table 5.4. Optimal lag combination

Var.#	Input Variables for Soft Sensor	Optimal Lag
1	Reactor 1 exit temperature (u1)	(k-1) to (k- 5)
2	FC403B.PV (u2)	(k-8) to (k- 9)
3	TR302-8.PV (Exit temperature of D-301) (u3)	-
4	FI403A.PV (Flow rate into reactor 1) (u4)	(k-6) to (k-10)
5	TI448.PV (Reactor 1 inlet temperature) (u5)	-
6	Inlet acetylene (u6)	k

5.5.3 Regression Coefficients and its Performance

The optimal regression coefficients obtained from the above procedure are compiled in Table 5.5. Some relative and absolute errors for validation data sets were calculated using the obtained regression coefficients and the performance of the resulting model is shown in Table 5.6.

Table 5.5. Optimal regression coefficients

Var.#	Input Variables for Soft Sensor	Lag	Regression Coefficient
1	Reactor 1 exit temperature	(k-1)	-5.2791
2	Reactor 1 exit temperature	(k-2)	-9.003
3	Reactor 1 exit temperature	(k-3)	-29.301
4	Reactor 1 exit temperature	(k-4)	-51.227
5	Reactor 1 exit temperature	(k-5)	-32.73
6	FC403B.PV	(k-8)	-8.9369
7	FC403B.PV	(k-9)	-11.745
8	FI403A.PV (Flow rate into reactor 1)	(k-6)	3.9173
9	FI403A.PV (Flow rate into reactor 1)	(k-7)	3.1845
10	FI403A.PV (Flow rate into reactor 1)	(k-8)	6.2039
11	FI403A.PV (Flow rate into reactor 1)	(k-9)	4.6136
12	FI403A.PV (Flow rate into reactor 1)	(k-10)	5.7618
13	Inlet acetylene	k	5457.7
14	Constant Term	***	0.40661

Table 5.6. Performance summary

Data Set #	MAE	MDAE	MINAE	MAXAE	count2	count5
1	14.1415	11.2109	0.0451	217.5522	0.9580	0.9991
2	12.7692	10.8101	0.0158	79.1467	0.9712	1
3	11.9596	9.9678	0.1334	55.2722	0.9693	1
5	14.5438	12.2813	0.0030	82.5523	0.9671	1

MAE = mean absolute error

MDAE = median absolute error

MINAE = minimum absolute error

MAXAE = maximum absolute error

count2 = 2% relative error

count5 = 5% relative error

The identified model almost always gives predictions that are less than 5% in relative error and its predictions are within 2% relative error 95% of the time. The quality of the model is indeed very good.

5.5.4 Validation on Other Data Sets

The validation of the model obtained from DSAR method using SET4 was performed on other four data sets. These validation figures are shown in Figures 5.33 to 5.36. The top subplot is the validation for all data points of the certain data set, the left and center subplots (bottom) show zoomed versions of the top plot, and the right bottom

subplot shows the scatter plot between model prediction and actual measurement (the X-axis represents the measured data and Y-axis represents the model output). In the scatter plot, if most points lie on the diagonal line one can conclude that a good agreement exists between model output and measured plant data. These validation plots do indicate the adequacy and usefulness of the model. The intersample predictions made by the model (shown as continuous line) on the four validation data sets are shown in Figure 5.37 through Figure 5.40. The constant term in the model is updated at every time a new measurement (indicated by '*') comes in.

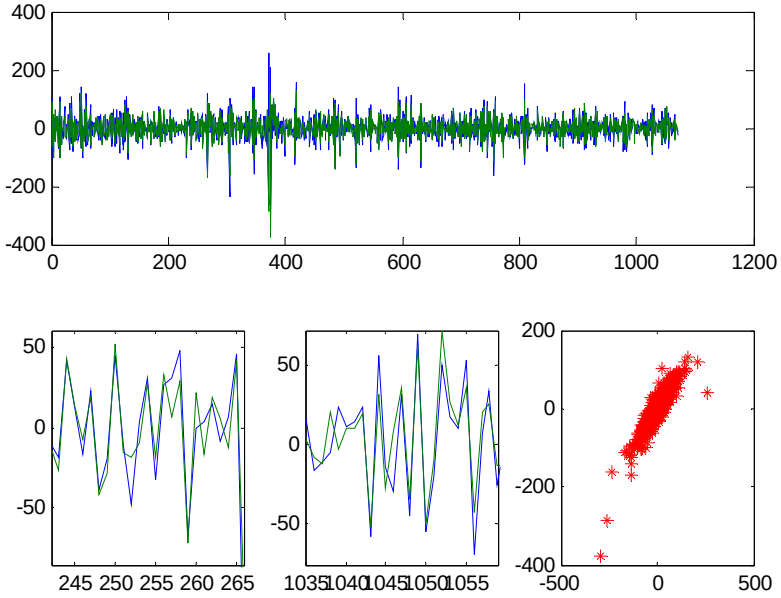

Figure 5.33: Validation on SET 1

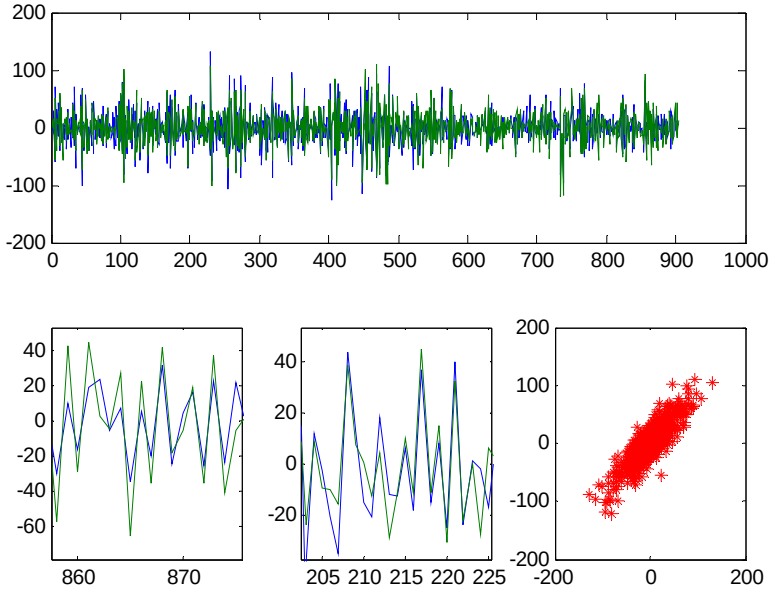

Figure 5.34: Validation on SET 2

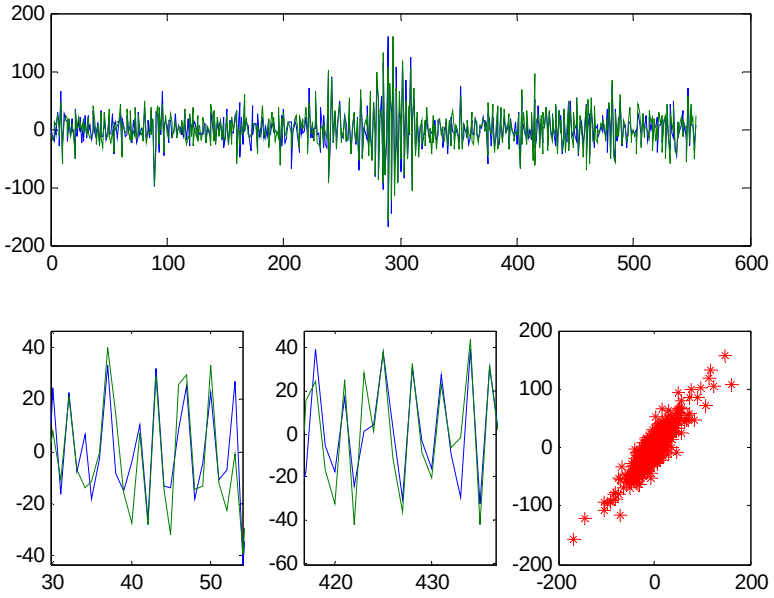

Figure 5.35: Validation on SET 3

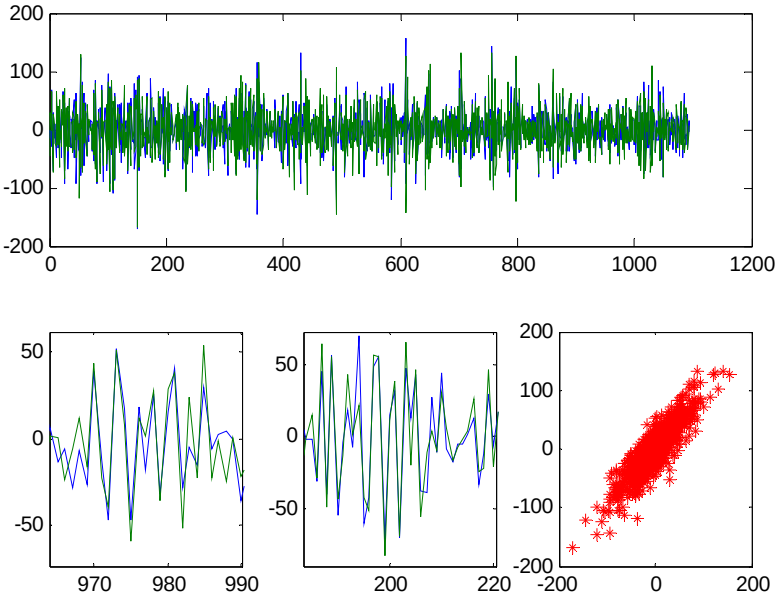

Figure 5.36: Validation on SET 5

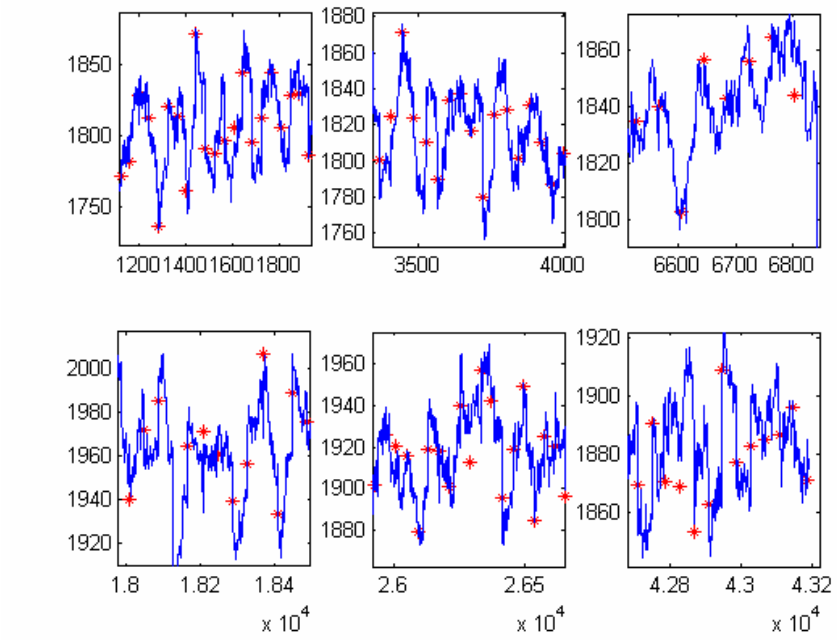

Figure 5.37: Validation on SET 1

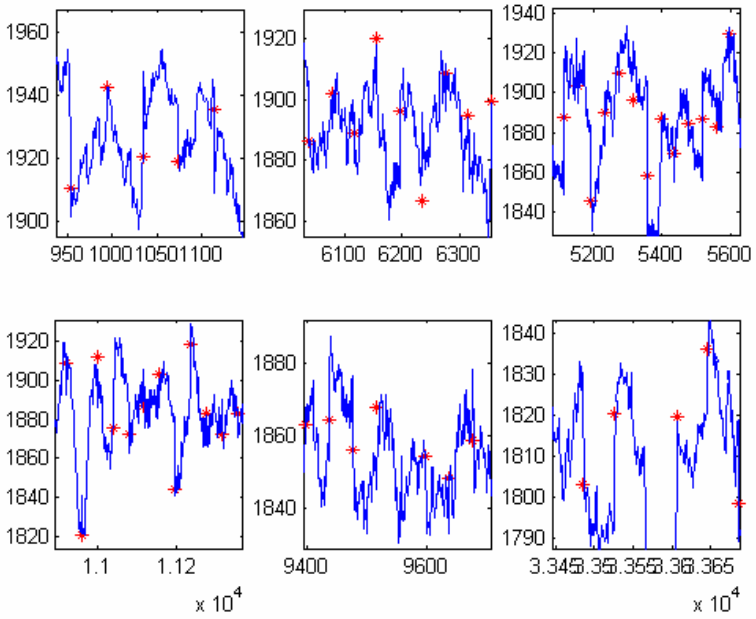

Figure 5.38: Validation on SET 2

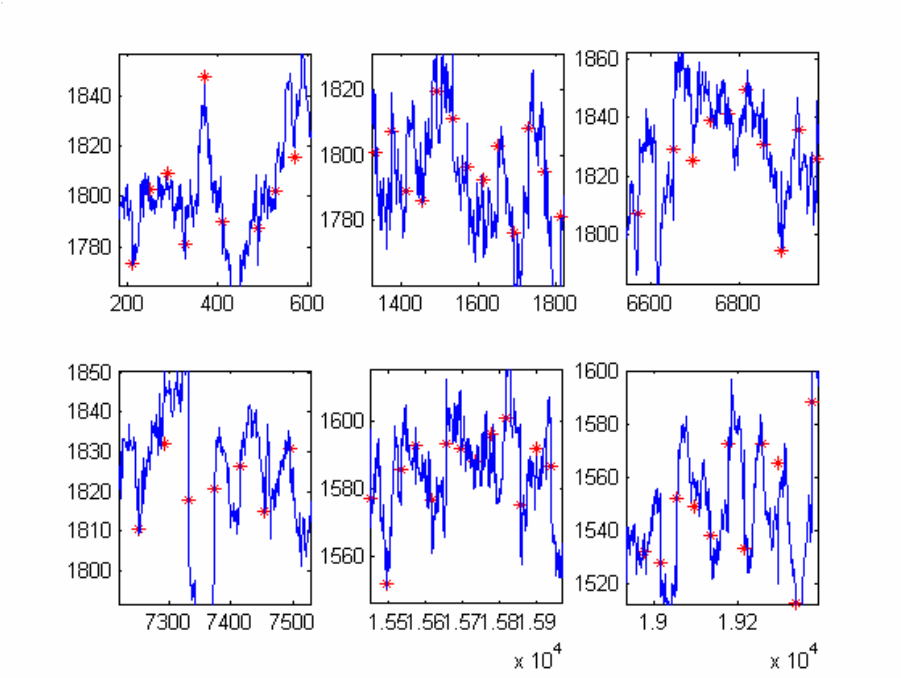

Figure 5.39: Validation on SET 3

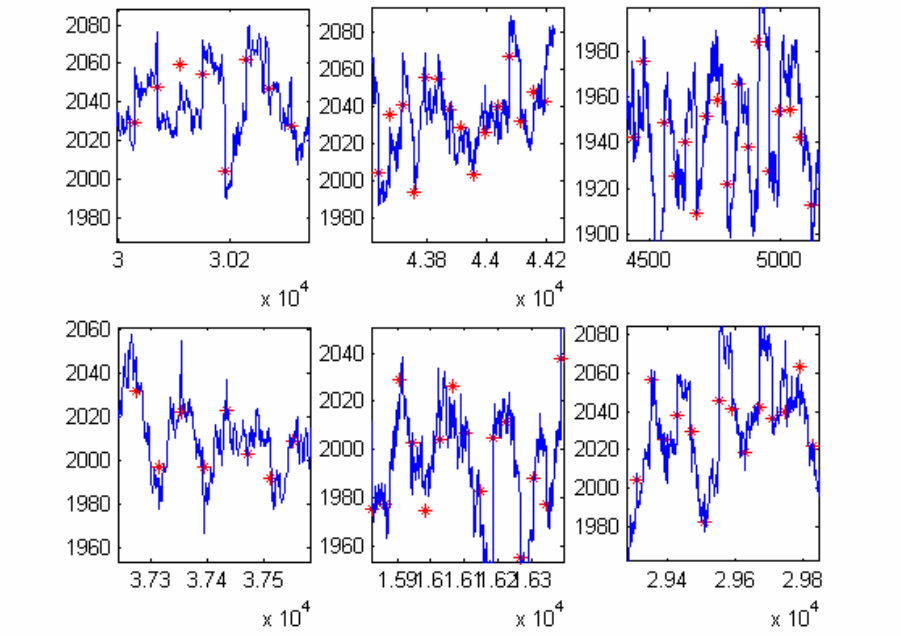

Figure 5.40: Validation on SET 5

5.6 Conclusions

The effect of input signal and the effect of γ value on the identified methods are evaluated with simulated and experimental data set for both linear and nonlinear multirate systems. Also, exploration of a good input signal for the nonlinear system identification for Hammerstein and Wiener type multirate system is done. The usefulness of DSAR identification method for modeling industrial data sets is practically confirmed by following the strategies mentioned in chapter 4. From these studies, it is seen that the input signal has a significant impact on the identification of both linear and nonlinear multirate systems.

CHAPTER 6

CONCLUSIONS

6.1 Contributions of the Thesis

The lifting technique is a commonly used approach for the identification of process from multirate data; lifting operator is used to convert the single input single output multirate identification problem into a single-rate multivariable identification problem. Subspace Identification algorithms such as SubSpace based State Space identification (N4SID), canonical variate analysis (CVA), and multivariable output error state space (MOESP) methods are used to identify the slow-rate linear dynamic model from which the fast-rate model is obtained. Among the available methods to construct the fast rate model from the slow rate model, three methods - the eigenvalue method, matrix roots approach and a model reduction approach have been employed. The model reduction based approach is a contribution of this work. The effect of sampling frequency and the different kinds of input signal on this technique are explored using simulation examples.

Both experimental and simulated data sets have been employed to identify linear and nonlinear models from multirate data by using lifting technique. Various ratios of sampling intervals (γ = 2 to 5) were considered. From our observation, the value of γ affects the identification result and the modified alternate approach (model reduction based approach) is explored as a remedy for this problem. In this work, γ = 1 is studied as the special case of multirate system to measure the usefulness of developed algorithms to the single-rate system identification. This thesis work has resulted in methods for identification of Hammerstein model and Wiener model from multirate

data. The developed methods are applicable to SISO and MISO multirate nonlinear systems. The random input signal is proposed as a best excitation signal for the identification of nonlinear multirate system especially for Hammerstein and Wiener type nonlinear systems (the evaluation is provided with simulated case studies).

DSAR method is compatible for the large γ values and its application to industrial data set is explored. Initially, the measurement of process inputs are available at every one minute intervals and process output is available at every 40 minute intervals. This is really unhelpful for the operator and can lead to poor process operation. Using DSAR, a soft sensor which can predict the output acetylene concentration for every minute is developed. The effect of different kind of input signals on DSAR method identification and remedy for the situation in which PRBS is employed as a input signal are also provided. The "extended" DSAR method is explored for non-integer ratio of sampling interval (non-integer γ) in which the fast-rate model with based sampling interval p is extracted from input sampling interval $2p$ and output sampling interval $3p$ (the evaluation with experimental data set is also provided).

6.2 Future Work

DSAR method is useful for industrial data multirate system identification in which the ratio of sampling intervals is very large. In this work, it is applied to linear systems only. One could attempt to solve the nonlinear multirate system identification problem using the DSAR method. This would extend the practical utility of DSAR.

The models developed here could be used for designing controllers and the effect on the control loop performance can be studied.

REFERENCES

Billings, S. A., and W. S. F. Voon. Correlation based model validity tests for nonlinear models, Int. J. Control, v. 44, pp. 803-822. 1986.

Boyd, S., and L. O. Chua. Fading memory and the problem of approximating nonlinear operators with Volterra series, IEEE Trans. Circuits and Systems, 32(11), pp. 1150-1161. 1985.

Bruls, J., C. T. Chou, B. R. J. Haverkamp and M. Verhaegen. Linear and Non-linear System Identification using Separable Least-Squares, European Journal of Control, (1999)5: 116-128. 1999.

Chen, T. and B. Francis. Optimal Sampled-data Control Systems, Springer, London. 1995.

Dayal, B. S. and J. F. MacGregor. Identification of Finite Impulse Response Models: Methods and Robustness Issues, Ind. Eng. Chem. Res., 35, pp.4078-4090. 1996.

Edgar, T. F., D. M. Himmelblau, and L. S. Lasdon. Optimization of Chemical Processes. 2nd ed. pp. 327, McGraw-Hill, New York, 2001.

Eskinat, E., S. H. Johnson and W. L. Luyben. Use of Hammerstein Models in Identification of Nonlinear Systems, AIChE Journal, 37(2), 255-268. 1991.

Friedland, B. Sampled-data control systems containing periodically varying members, In Proc. First IFAC Congress, pp. 361-367. 1960.

Gerald, C. F. and P. O. Wheatley. Applied Numerical Analysis, 4th ed. Addison-Wesley Publishing Company. 1989.

Gopaluni, R. B., H. Raghavan, S. L. Shah. System Identification from Multi-rate Data, Presented at the International Symposium on Advanced Control of Chemical Processes (ADCHEM), Hong Kong, June 2003.

Hagenblad, A. Aspects of the Identification of Wiener Models, Linköping Studies in Science and Technology Thesis, Linköping. 1999.

Henson, M. A. and D. E. Seborg. Adaptive nonlinear control of a pH neutralization process. IEEE Trans. Control System Technology, 2(3), 169-182. 1994.

Henson, M. A. and D. E. Seborg. (ed). Nonlinear Process Control. Prentice Hall PTR. 1997.

Hunter, I. W., and M. J. Korenberg. The identification of nonlinear biological systems: Wiener and Hammerstein cascade models, Biol. Cybern, 55, pp. 135-144. 1986.

Khargonekar, P. P., K. Poolla, and A. Tannenbaum. Robust Control of Linear Time-Invariant Plants using Periodic Compensation, IEEE Trans. on Automatic Control, AC-30(11),1088-1096. 1985.

Kranc, G. M. Input-output analysis of multirate feedback systems, IEEE Trans. on Automatic Control, 30, pp. 21-28. 1957.

Kresta, J., J. F. MacGregor, and T. E. Marlin, Multivariate Statistical Monitoring of Process Operating Performance, Can. J. Chem. Eng., 69, 35-47, 1991.

Lakshminarayanan, S. Process Characterization and Control using Multivariate Statistical Techniques. Ph.D Thesis, University of Alberta. 1997.

Lakshminarayanan, S. Canonical Variate Analysis: A Primer. Internal Report, Mitsubishi Chemical Corporation, 1997.

Lakshminarayanan, S. A data selection and regression method for identifying Multirate system: Application to an industrial acetylene reactor, Internal Report, Mitsubishi Chemical Corporation, 2000.

Larimore, W. E. Canonical Variate Analysis in Identification, Filtering and Adaptive Control, In: Proceedings of the 29th IEEE Conference on Decision and Control, pp. 596-604, 1990.

Levine, W. S. The Control Handbook. pp. 1033-1054, New York: CRC Press, IEEE Press. 1996.

Li, D. System Identification and Control of Multirate Systems. Ph.D Thesis, University of Alberta. 2001.

Li, D., S. L. Shah and T. Chen. Identification of fast-rate models from Multirate data, Int. J. Control, 74(7), 680-689. 2001.

Li, D., S. L. Shah, T. Chen, K. Z. Qi. Application of dual-rate modeling to CCR octane quality inferential control, to appear in IEEE Trans. On Control Systems Tech. 2002.

Ljung, L. System identification: Theory for the User, 2nd ed. Prentice Hall, Englewood Cliffs, N.J. 1999.

Ljung, L. The System Identification Toolbox: The Manual, 5th ed. The Math Works Inc, Natick, MA, 2001.

MacGregor, J. F., T. Kourti, J. V. Kresta. Multivariate Identification: A Study of Several Methods, IFAC ADCHEM Conference Proceedings, Tolouse, France, Oct 1991.

Moonen, M., B. De Moor, L. Vandenberghe and J. Vandewalle. On- and Off-line identification of linear state space models, Int. J. Control, 49, pp.219-232. 1989.

Narendra, K. S., and P. G. Gallman. An Iterative Method for the Identification of Nonlinear Systems using a Hammerstein Model, IEEE Trans. on Automatic Control, pp.546-550. 1966.

Nikolaou, M. and P. Vuthandam. FIR Model Identification: Achieving Parsimony through Kernel Compression with Wavelets, Submitted to AIChE, March 1997.

Pawlak, M. On the series expansion approach to the identification of Hammerstein systems. IEEE Trans. on Automatic Control, 36, pp.763-767. 1991.

Söderström, T. and P. Stoica. System Identification, Englewood Cliffs, NJ: Prentice Hall. 1989.

Stoica, P., and T. Söderström. Instrumental Variable Methods for Identification of Hammerstein Systems. Int. J. Control, 35, 459. 1982.

The Math Works, Inc. Control System Toolbox User's Guide (Version 4). Jan 1998.

Van Overschee, P. and B. De Moor. Subspace algorithms for the stochastic identification problem. 30th IEEE Conference on Decision and Control, Brighton, U.K., 1321-1326. 1991a.

Van Overschee, P. and B. De Moor. Subspace algorithms for the Stochastic Identification Problem, ESAT/SISTA report 1991-26, Kath. Universiteit Leuven, Dept. E. E., Belgium. Automatica, 29, 649-660. 1991b.

Van Overschee, P. and B. De Moor. N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic System, Automatica, 30(1), pp.75-93. 1994.

Verhaegen, M. and P. Dewilde. Subspace model identification Part 1. The output-error state-space model identification class of algorithms, Int. J. Control, 56(5), 1187-1210. 1992a.

Verhaegen, M. and P. Dewilde. Subspace model identification Part 2. Analysis of the elementary output-error state-space model identification algorithm, Int. J. Control, 56(5), 1211-1241. 1992b.

Verhaegen, M. Subspace model identification Part 3. Analysis of the elementary output-error state-space model identification algorithm, Int. J. Control, 58(3), 555-586. 1993.

Verhaegen, M. Identification of the Deterministic Part of MIMO State Space Models given in Innovations Form from Input-Output data, Automatica, 30(1), 61-74. 1994.

Verhaegen, M. and X. Yu. A class of subspace model identification algorithms to identify periodically and arbitrarily time-varying systems, Automatica, 32, 201-216. 1995.

Verhaegen, M. and D. Westwick. Identifying MIMO Hammerstein systems in the context of subspace model identification methods, Int. J. Control, 63(2), 331-349. 1996.

Wang, J., T. Chen, B. Huang. Multirate sampled-data systems: computing fast-rate models, Journal of Process Control, 14, 79-88. 2004.

Westwick, D. and M. Verhaegen. Identifying MIMO Wiener systems using subspace model identification methods, Signal Processing, 52, 235-258. 1996.

Westwick, D. T. and R. E. Kearney. Separable Least Squares Identification of Nonlinear Hammerstein Models: Application to Stretch Reflex Dynamics, Annals of Biomedical Engineering, vol. 29(8), pp. 707-718. 2001.

Zadeh, L. A. From circuit theory to system theory, Proc. IRE, v. 50, pp. 856-865. 1962.

Zhu, Y. Multivariable System Identification for Process Control, 1st ed. Elsevier Science, UK. 2001.

BIOGRAPHY

May Su Tun was born in Henzada, the second capital of Ayeyarwadi Division, Myanmar. She grew up in Pathein (capital of Ayeyarwadi Division; near Yangon, the capital of Myanmar) after her parents moved there when she was 12. She finished her high school education in 1990 and joined the Industrial Chemistry Department, Yangon University in 1993. She got her first degree (B.Sc. (Honors)) in 1996. Her interest in the chemical industry and engineering encouraged her to join the Chemical Engineering Department at Yangon Technological University from where she earned her second Bachelor degree, B.Eng. (Chemical) in 2001. She joined the Chemical and Biomolecular Engineering Department, National University of Singapore in July, 2002.