Université d'Ottawa · University of Ottawa

Université d'Ottawa · University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

BAI, Shuanghua

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

M.A.Sc. (Chemical Engineering)

GRADE - DEGREE

Department of Chemical Engineering

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

Assessment of Controller Performance with Embedded Data Reconciliation

J. Thibeault

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

D. McLean

CO-DIRECTEUR DE LA THÈSE - THESIS CO-SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

M. A. Dubé

D. G. Taylor

J.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

SIGNATURE

DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

Assessment of Controller Performance with Embedded Data Reconciliation

by

Shuanghua Bai

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfilment of the requirements for the degree of

Master of Applied Science

in the

Department of Chemical Engineering
University of Ottawa
May 2003

Your file Votre référence
ISBN: 0-612-90023-1
Our file Notre référence
ISBN: 0-612-90023-1

Statement of Contributions of Collaborators

I hereby declare that this thesis is accomplished under the supervision of Professors David D. McLean and Jules Thibault of the Department of Chemical Engineering at the University of Ottawa. The material in this thesis has been the outcome of two years of research efforts by the author and his two supervisors. I personally wrote the thesis and my supervisors made editorial corrections.

Signature: _____ Date: _____

Abstract

Process measurements contain some degree of error that can be random or systematic. Random measurement noise is usually of high frequency, and results in high-frequency oscillations of manipulated variables that deteriorate the performance of the control system. Data reconciliation is a procedure that makes use of process models along with process measurements to give more precise and consistent estimates of process variables. Data reconciliation has been traditionally used to provide a more representative set of data to calculate steady-state inventories and process yields. For dynamic systems, the use of data reconciliation is relatively nascent. This work examined the potential use of data reconciliation in closed-loop control as a filter to attenuate the noise in measurements of the controlled variables so that the controllers can act on less variable, more accurate inputs. Data reconciliation filters were implemented in simulations of a PID control system for a binary distillation column. Results showed that data reconciliation could efficiently reduce the propagation of measurement noise in control loops, so that the overall performance of the controller was enhanced.

The presence of measurement noise usually results in detuned controllers in order to prevent excessive high-frequency variations of manipulated variables. However, with detuned controllers, the process is characterized by sluggish dynamic responses. In this work, a Dynamic Data Reconciliation (DDR) algorithm, embedded within the structure of feedback control loops, was developed to reconcile noisy measurements. The developed DDR algorithm employed discrete dynamic models, both phenomenological

and empirical. Models were not assumed to be exact constraints on the process variables. Instead, the data reconciliation algorithm was based on a compromise between the actual measurements and the model predicted values. Simulation results for a storage tank and a distillation column under PI control demonstrated that use of the DDR algorithm within the structure of a feedback control loop not only resulted in faster responses of controlled variables, but also reduced variations of manipulated variables.

In this thesis, contributions were made in three areas: investigation of the impact of measurement noise on the performance of controllers, formulations of data reconciliation filters and their relevant applications, and development of dynamic data reconciliation strategies and their implementations.

Résumé

Les mesures des variables d'un procédé sont souvent entachées d'erreurs qui peuvent être aléatoires ou systématiques. Le bruit aléatoire de mesure est habituellement de haute fréquence, et produit des oscillations à haute fréquence des variables manipulées qui détériorent les performances du système de commande. La méthode de réconciliation des données est une technique qui utilise des modèles décrivant le procédé et des mesures des variables de procédé pour obtenir des évaluations plus précises et plus représentatives des variables réelles de procédé. La réconciliation des données a été traditionnellement employée pour fournir un ensemble plus représentatif des données permettant d'effectuer des inventaires et des calculs de rendement. Pour les systèmes dynamiques, l'utilisation de la réconciliation des données est relativement naissante. Ce travail a examiné l'utilisation de la réconciliation des données dans la commande en circuit fermé comme filtre pour atténuer le bruit dans les mesures des variables régulées de sorte que les contrôleurs puissent agir sur des entrées plus stables et plus précises. Des filtres de réconciliation des données ont été mis en application dans les simulations d'un système de commande PID pour une colonne binaire de distillation. Les résultats ont montré que la réconciliation des données pouvait réduire la propagation du bruit de mesure dans des boucles d'asservissement, de sorte que la performance globale du contrôleur est améliorée.

La présence du bruit dans les mesures exige que la vitesse de réponse des contrôleurs soit réduite afin d'empêcher les variations excessives des variables manipulées. Cependant,

avec les contrôleurs moins agressifs, le procédé est caractérisé par des réponses dynamiques plus lentes. Dans ce travail, un algorithme dynamique de la réconciliation des données (DDR), incorporé dans la structure même des boucles de rétroaction, a été développé pour réconcilier les mesures bruitées. L'algorithme développé a utilisé des modèles dynamiques discrets de nature phénoménologique et empirique. Ces modèles ne sont pas des représentations exactes pouvant prédire les variables de procédé. En effet, l'algorithme de réconciliation des données a été basé sur un compromis entre les mesures réelles et les valeurs prédites par les modèles. Les résultats de simulation pour un réservoir d'entreposage et une colonne de distillation, utilisant des régulateurs PI, ont démontré que l'utilisation de l'algorithme DDR dans la structure d'une boucle rétroaction a non seulement eu comme conséquence des réponses plus rapides des variables commandées, mais des variations également réduites des variables manipulées.

Dans cette thèse, des contributions ont été apportées dans trois aspects de la commande automatique: une étude sur l'impact du bruit de mesure sur l'exécution des contrôleurs, les formulations des filtres de réconciliation des données et de leurs applications.

Acknowledgements

I would like to thank my supervisors, Dr. David D. McLean and Dr. Jules Thibault, for giving me the opportunity to carry out this interesting research topic under their enthusiastic guidance and support. I also wish to express my gratitude to them for helping me to improve the language of the thesis. This thesis would not have been possible without their fruitful instruction and comments throughout this work.

I am very grateful to Professor Marc A. Dubé, David Taylor and Arturo Macchi, who have thoroughly read this thesis and provided invaluable advice. I also would like to thank everyone in the Department for their companionship and help during the past two years.

Last, but not least, I thank my family for their endless support during my pursuit of this degree.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Reliable and accurate process data are paramount for process monitoring, optimization, and control. Unfortunately, plant measurements contain errors that often invalidate, or deteriorate the performance of optimization and control strategies. Data reconciliation is a technique to improve the accuracy and precision of measurements by reducing the effect of errors in the data. It uses measured data and process models to provide more accurate and precise estimates of process variables. More importantly, the reconciled data are consistent with the known relationships between process variables, such as mass and heat balances.

The interest in applying data reconciliation techniques started in the 1980s when plant management realized the benefits of having access to more reliable estimates of process data. Previously, raw, unconditioned data were used for process modeling and optimization, but the results could be unrelaible, since erroneous data were used to represent the state of the process. Research and development during the past 30 years have led to two major types of applications for data reconciliation in various chemical, petrochemical, and mineral processing industries. Mass and heat balance reconciliation for a plant was the first type of application. The simplest example was the off-line reconciliation of plant flow rates around a process unit. The reconciled flow rates satisfy

the overall mass balance of the unit. Parameter estimation is another important application for data reconciliation that has been developed. Accurate, precise estimates of model parameters, such as heat transfer coefficients, fouling factors of heat exchangers and distillation tray efficiencies are required in order to obtain reliable model predictions for process simulation, design and optimization. One approach to the parameter estimation is to solve the estimation problem simultaneously with the data reconciliation problem. The model parameters are treated as unmeasured variables or tuning parameters that are adjusted via the data reconciliation algorithm to match the plant measurements and process model constraints. The reconciled model parameters are expected to be more accurate and can be used with greater confidence.

Although, these two types of applications for data reconciliation could be implemented on-line via a data historian provided by a distributed control system, the two applications have been limited to steady-state processes where averaged data (e.g., over a one-hour period) have been used. Steady-state data reconciliation is satisfactory for applications such as steady-state simulation, or on-line optimization where optimal set points are calculated once every hour. However, disturbances occur frequently in chemical processes, and most chemical processes are intrinsically dynamic. Dynamic data reconciliation strategies need to be developed for dynamic processes. Moreover, in regulatory control, measurements are usually corrupted by process noise. When the noisy measurements are used by controllers, controller performance deteriorates. Application of

a dynamic data reconciliation algorithm to process measurements allows controllers to access more accurate and precise data at every sampling instant.

Dynamic data reconciliation is a relatively nascent technology compared with steady-state data reconciliation. Most of the previous work dealing with dynamic data reconciliation focussed on the problems of process monitoring, inventory calculations, and process state estimation. Only a few investigations addressed the problem of applying dynamic data reconciliation to process control.

The prime objectives of this thesis were to develop and test dynamic data reconciliation algorithms embedded in the structure of standard PID control loops. Of particular concern in this study was the effectiveness of data reconciliation in overcoming the degradation of controller performance in the presence of significant measurement noise.

To meet these objectives, dynamic simulations of feedback systems for storage tank level control and for binary (benzene-toluene) distillation control were carried out. The development of the simulators followed conventional steps including problem definition, mathematical modeling, selection of numerical algorithm, program coding and simulation validation. The resulting simulators provide much flexibility for this and future studies, because they are not only applicable under standard PID control, but can be easily modified to handle other complex control strategies. The simulators were coded in FORTRAN. Methodologies pertaining to the dynamic simulation of a distillation column

are presented in Appendix A. To validate the simulator, simulations of open-loop and closed-loop scenarios were performed. The simulation results are discussed in Appendix B. Appendix C presents the approaches employed to identify empirical dynamic models used in the dynamic data reconciliation for the control of the distillation column. Appendix D lists basic physical properties of benzene and toluene that were used in the simulations.

The core of this thesis essentially consists of two papers. Preceding these two papaers, Chapter 2 presents background information and a review of the literature for both steady-state and dynamic data reconciliation in terms of their formulations, algorithms, and applications. Chapter 3, the first paper, develops quasi-steady-state data reconciliation algorithms that can be used as filters within the control structure for the binary distillation column. The effect of measurement noise on the performance of the controllers is investigated, and the performance of the data reconciliation filters is evaluated. Chapter 4, the second paper, presents a dynamic data reconciliation strategy embedded inside the control loops for filtering measurement noise, for both simple level control in a storage tank and multivariable control of the binary distillation column. The main conclusions and recommendations are summarized in Chapter 5.

CHAPTER 2

Background & Literature Review

2.1 Basic Concepts in Data Reconciliation

Measured process data inevitably contain some inaccurate information since measurements are obtained with imperfect instruments, which have their own accuracy. In addition, signal transmission, power fluctuation, improper instrument installation and miscalibration are other sources of errors. The raw process measurements can be expressed as

$$\mathbf{y} = \mathbf{x} + \varepsilon \tag{2.1}$$

where $\mathbf{y}$ is a $M \times 1$ vector of measurements for M process variables, $\mathbf{x}$ is a $M \times 1$ vector of true values for the M process variables and ε is a $M \times 1$ vector of measurement errors.

The measurement errors can be classified as random errors or gross errors. Random errors are usually attributed to irreproducible measurement noise that is generally assumed to be normally distributed with a mean value of zero and known variance. On the other hand, gross errors occur when the measurement device fails, when nonrandom events affect the process, and when instruments provide a systematic bias that is consistently erroneous, either higher or lower than true value of the process variable.

When the corrupted measurements are used for state estimation and process control, the state of the process is misrepresented, control performance may deteriorate, and in some cases leading to unsafe operating conditions. Normally the impact of gross errors is more significant than that of small random errors. To improve the accuracy and precision of the estimated values of process variables, the techniques of data reconciliation (DR) and gross error detection have been developed (Narasimhan et al., 2000; Romagnoli et al., 2000). Although it should be noted that gross error detection is an important step in processing the raw measurements, the present study is only concerned with data reconciliation.

Data reconciliation is the estimation of process variables based on information contained in process measurements and process models to arrive at estimates that are more precise and accurate than estimates based on only measurements or models. In general, the optimal estimates by DR are solutions to a constrained least-squares or maximum likelihood objective function, where the constraints are models representing the process. With the assumption of normally distributed measurements, a least-squares objective function is conventionally formulated for the data reconciliation problem as

$$\begin{aligned} &\text{Minimize} && J(\hat{\mathbf{y}},\hat{\mathbf{z}}) = (\mathbf{y}-\hat{\mathbf{y}})^T \mathbf{V}^{-1}(\mathbf{y}-\hat{\mathbf{y}}) \\ &\text{subject to} && \mathbf{f}(\hat{\mathbf{y}},\hat{\mathbf{z}}) = \mathbf{0} \\ & && \mathbf{g}(\hat{\mathbf{y}},\hat{\mathbf{z}}) \geq \mathbf{0} \end{aligned} \tag{2.2}$$

where $\hat{\mathbf{y}}$ is an $M \times 1$ vector of estimates (reconciled values) for the M raw measurements, $\hat{\mathbf{z}}$ is a $N \times 1$ vector of estimates (reconciled values) for unmeasured process variables, $\mathbf{V}$ is an $M \times M$ covariance matrix of the measurements, $\mathbf{f}$ is a $C \times 1$ vector describing the functional form of model equality constraints, and $\mathbf{g}$ is a $D \times 1$ vector describing the functional form of model inequality constraints that includes simple upper and lower bounds.

A simple illustrative example of a data reconciliation problem is presented in Figure 2.1. A cooling-water station provides circulation water for four plants. All the flow rates of the circulation water are measured in this network. The true values of the flow rates, the raw measurements and their standard deviations are listed in Table 2.1.

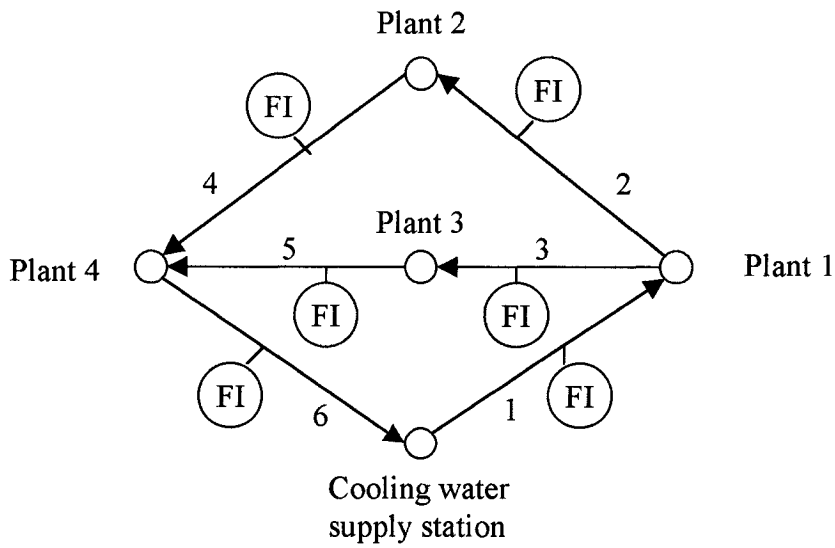

Figure 2.1 An example of reconciling measured flow rates for a circulation cooling-water network.

Table 2.1 Data reconciliation for a circulation cooling water network

Stream No.	True (kt/h)	Raw (kt/h)	Standard deviation (kt/h)	Reconciled (kt/h)
1	100	101.91	0.82	100.32
2	64	64.45	0.53	64.61
3	36	34.65	0.46	35.71
4	64	64.20	0.71	64.61
5	36	36.44	0.45	35.71
6	100	98.88	1.20	100.32

Applying the data reconciliation algorithm formulated by equation (2.2), the vector of the raw measurements for the circulation water flow can be written as

$$\mathbf{y} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = \begin{bmatrix} 101.91 \\ 64.45 \\ 34.65 \\ 64.20 \\ 36.44 \\ 98.88 \end{bmatrix}.$$

Assuming the six measurements are not correlated, the covariance matrix of the six measurements, **V**, is in diagonal form, and it can be written as

$$\mathbf{V} = \begin{bmatrix} 0.6724 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.2809 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2116 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5041 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.2025 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.44 \end{bmatrix}.$$

The process model constraints in this case are the mass balances for each plant (node) in the network. This is to say that the reconciled values should satisfy the mass balances at each node

Plant 1: $\hat{F}_1 - \hat{F}_2 - \hat{F}_3 = 0$

Plant 2: $\hat{F}_2 - \hat{F}_4 = 0$

Plant 3: $\hat{F}_3 - \hat{F}_5 = 0$

Plant 4: $\hat{F}_4 + \hat{F}_5 - \hat{F}_6 = 0$.

Writing the process model constraints above in a compact form gives $\mathbf{A}\hat{\mathbf{y}} = \mathbf{0}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}, \text{ and } \hat{\mathbf{y}} = \begin{bmatrix} \hat{F}_1 \\ \hat{F}_2 \\ \hat{F}_3 \\ \hat{F}_4 \\ \hat{F}_5 \\ \hat{F}_6 \end{bmatrix}.$$

The matrix **A** is called the *incidence matrix*. It should be noted that in the incidence matrix **A**, each row represents each node and each column represents each flow, respectively. Each element in **A** is either +1, -1 or 0 depending on whether the corresponding flow is an input stream, an output stream, or not associated with this node.

Therefore, the data reconciliation problem for the circulation water network becomes

$$\text{minimizing} \quad J(\hat{\mathbf{y}}) = (\mathbf{y} - \hat{\mathbf{y}})^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\mathbf{y}}) \tag{2.3}$$

$$\text{subject to} \quad \mathbf{A}\hat{\mathbf{y}} = \mathbf{0}$$

The optimization problem of (2.3) can be solved using Lagrange multipliers (Romagnoli et al., 2000). The reconciled flow rates are obtained by

$$\text{Minimizing} \quad J(\hat{\mathbf{y}}) = (\mathbf{y} - \hat{\mathbf{y}})^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\mathbf{y}}) - 2\lambda^{\mathrm{T}} \mathbf{A}\hat{\mathbf{y}} \tag{2.4}$$

where $\lambda^{\mathrm{T}} = [\lambda_1 \quad \lambda_2 \quad \lambda_3 \quad \lambda_4]$. The necessary conditions of the minimization of (2.4) are

$$\frac{\partial \mathrm{J}}{\partial \hat{\mathbf{y}}} = -2\mathbf{V}^{-1}(\mathbf{y} - \hat{\mathbf{y}}) - 2\mathbf{A}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{0} \tag{2.5}$$

$$\frac{\partial \mathrm{J}}{\partial \boldsymbol{\lambda}} = \mathbf{A}\hat{\mathbf{y}} = \mathbf{0}$$

Premultiplying each term by the covariance matrix, **V**, in (2.5) yields

$$\mathbf{y} - \hat{\mathbf{y}} + \mathbf{V}\mathbf{A}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{0} \tag{2.6}$$

Premultiplying each term by the incidence matrix, **A**, in (2.6), and applying $\mathbf{A}\hat{\mathbf{y}} = \mathbf{0}$ yield

$$\mathbf{A}\mathbf{y} + \mathbf{A}\mathbf{V}\mathbf{A}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{0} \tag{2.7}$$

Rearranging equation (2.7) gives

$$\boldsymbol{\lambda} = -(\mathbf{A}\mathbf{V}^{-1}\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{A}\mathbf{y} \tag{2.8}$$

Substituting $\boldsymbol{\lambda}$ in (2.6) and rearranging the equation give the vector of reconciled data

$$\hat{\mathbf{y}} = \mathbf{y} - \mathbf{V}\mathbf{A}^{T}(\mathbf{A}\mathbf{V}\mathbf{A}^{T})^{-1}\mathbf{A}\mathbf{y} \tag{2.9}$$

Putting the values of $\mathbf{y}$, $\mathbf{V}$ and $\mathbf{A}$ into (2.9) results in the reconciled values for each flow rate in the circulation water network. The reconciled values for each measurement are also listed in Table 2.1. It is clear that the reconciled values satisfy the mass balance at each node. If there is no any gross error such as leak of water, inappropriate calibration of measurement device, the reconciled data should be more consistent and reliable than the raw measurements.

It is also important to clarify some concepts in data reconciliation technologies. Measured variables are classified as *redundant* and *nonredundant*, whereas unmeasured variables are classified as *observable* and *nonobservable*. A redundant variable is a measured variable that can be estimated by other measured variables via process models, in addition to its measurement. A nonredundant variable is a measured variable that cannot be estimated other than by its own measurements. An observable variable is an unmeasured variable that can be estimated from measured variables through physical models, while a nonobservable variable is a variable for which no information is available. For the illustrative example shown in Figure 2.1, all six flows are redundant variables. However, if the flows 2, 4 and 6 were not measured, flow 1 becomes a measured nonredundant but flows 3 and 5 are still redundant variables. The unmeasured flows of streams 2, 4 and 6, in this case, are observable because their values can be estimated by mass balances around the nodes.

A measurement is *spatially redundant* if there are more than enough data to completely define the process at any instant in time. On the other hand, a measurement is *temporally redundant* if its past measurements can be used to estimate the current state. Data reconciliation utilizes measurement redundancies to improve the accuracy and consistency of process data. The redundancies arise from the fact that at least some information about the process is known and relates measurements to each other. Such information usually comes from process mass and heat balances, process kinetics, thermodynamics, or even empirical process input-output models.

2.2 Steady-State Data Reconciliation

2.2.1 Linear Steady-State Data Reconciliation

At steady state, process variables satisfy steady-state models that are equality or inequality algebraic equations. The simplest data reconciliation problem occurs in reconciling process flow rates in a plant as illustrated in the previous example of the circulation water supply network presented in Figure 2.1. The model constraints in this case are the overall mass balances at each node (process unit). If all the flows are measured, the linear steady-state data reconciliation problem is formulated as (2.3)

$$\begin{aligned} &\text{Minimize} && J(\hat{\mathbf{y}}) = (\mathbf{y} - \hat{\mathbf{y}})^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\mathbf{y}}) \\ &\text{subject to} && \mathbf{A}\hat{\mathbf{y}} = \mathbf{0} \end{aligned} \tag{2.10}$$

where $\mathbf{A}$ is a $C \times M$ incident matrix, and C is the total number of nodes in the process. The analytical solution of (2.10) is given as equation (2.9) as demonstrated. The reconciled flow rates, $\hat{\mathbf{y}}$, can be shown to be normally distributed and unbiased, $E(\hat{\mathbf{y}}) = \mathbf{x}$.

However, in chemical plants not all flows are measured due to physical or economical reasons. When unmeasured flows are involved in the model constraints of equation (2.10), in this case, the unmeasured flows can be estimated in data reconciliation. Partitioning the constraint equations in terms of measured and unmeasured variables

$$\mathbf{A}_y\hat{\mathbf{y}} + \mathbf{A}_z\hat{\mathbf{z}} = \mathbf{0} \tag{2.11}$$

where the columns of $\mathbf{A}_y$ correspond to the measured flows, and those of $\mathbf{A}_z$ correspond to the unmeasured flows. The dimensions of $\mathbf{A}_y$ and $\mathbf{A}_z$ are $C \times M$ and $C \times N$, respectively.

The solution to the data reconciliation problem with constraints posed by equation (2.11) are obtained by eliminating the unmeasured flows, $\hat{\mathbf{z}}$, by pre-multiplying both sides of the constraint equation by a *projection matrix* $\mathbf{P}$ $(C \times C)$ such that $\mathbf{PA}_z = \mathbf{0}$. The data reconciliation problem with unmeasured flows then becomes

$$\text{Minimize} \qquad J(\hat{\mathbf{y}}) = (\mathbf{y} - \hat{\mathbf{y}})^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\mathbf{y}}) \tag{2.12}$$

$$\text{subject to} \qquad \mathbf{P}\mathbf{A}_y \hat{\mathbf{y}} = \mathbf{0}$$

The solution to the optimization problem of (2.12) is given by equation (2.9) in which matrix **A** is replaced by matrix **PA**.

The projection matrix, **P**, can be obtained efficiently by **QR** factorization of matrix $\mathbf{A}_z$ (Narasimhan, 2000). After the redundant measured flows, $\mathbf{y}$, are reconciled, putting the values of $\hat{\mathbf{y}}$ into equation (2.11) yields C linear equations with N variables. Since $C \geq N$, a least-squares solution gives the observable unmeasured flows as

$$\hat{\mathbf{z}} = -(\mathbf{A}_z^T \mathbf{A}_z)^{-1} \mathbf{A}_z^T (\mathbf{A}_y \hat{\mathbf{y}}). \tag{2.13}$$

Equation (2.13) assumes that all the unmeasured flows are observable. If there exist unobservable variables, equation (2.13) can't be applied. The identification of unobservable flows in data reconciliation techniques can be found in Romagnoli et al. (2000).

2.2.2 Nonlinear Steady-State Data Reconciliation

If, for example, both flow rates and compositions of process streams are measured, the component mass balances are nonlinear in terms of the measured flow rates and

compositions. The estimation of the flow rates and compositions by data reconciliation at process steady state forms a nonlinear data reconciliation problem. Prior knowledge about the process such as thermodynamic relationships and other complex correlations can also be taken into considerations in estimation of process variables. Furthermore, inequality constraints can also be imposed on the estimates. For example, flow rates can't be negative, and molar or mass fractions should be between 0 and 1.

In the case of the nonlinear steady-state data reconciliation problem that is only subject to equality constraints, two approaches, the method of Lagrange multipliers and the method of successive linear data reconciliation, have been developed to solve this problem. The Lagrange multiplier method for the problem is given by

$$\text{Minimizing} \quad J(\hat{\mathbf{y}}, \hat{\mathbf{z}}, \lambda) = (\mathbf{y} - \hat{\mathbf{y}})^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\mathbf{y}}) + \lambda^{\mathrm{T}} \mathbf{f}(\hat{\mathbf{y}}, \hat{\mathbf{z}}) \tag{2.14}$$

where λ is a vector of penalty parameters. The solution to the optimization problem of (2.14) can be obtained by setting the partial derivatives of the objective function with respect to the variables $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$, and λ to zero. Then, the remaining nonlinear algebraic equations can be simultaneously obtained by a Newton - like solver. The method of successive linear data reconciliation, on the other hand, handles the nonlinear data reconciliation as a series of linear data reconciliation problems by linearizing the nonlinear constraints, $\mathbf{f}(\hat{\mathbf{y}}, \hat{\mathbf{z}}) = \mathbf{0}$. At each iteration, linear approximations to the constraints are obtained, and linear data reconciliation is applied to obtain the solutions.

The procedure is repeated until the norms $\|\hat{\mathbf{y}}_k - \hat{\mathbf{y}}_{k-1}\|$ and $\|\hat{\mathbf{z}}_k - \hat{\mathbf{z}}_{k-1}\|$ satisfy acceptable tolerances (Narasimhan et al., 2000).

When the nonlinear steady-state data reconciliation problem involves both equality and inequality constraints, the optimization problem (2.2) can be solved by general purpose nonlinear programming (NLP) techniques, such as sequential quadratic programming (SQP) and generalized reduced gradient (GRG) methods. The SQP technique solves a nonlinear optimization problem by successively solving a series of quadratic programming problems. At each iteration, a quadratic function approximates the objective function and the nonlinear constraints are linearized around the current estimates. The GRG method gives the solutions by solving a series of successive linear programming (LP) problems. At each iteration, a linear programing approximation is obtained by linearizing both the objective function and the constraints. Detailed descriptions of NLP techniques can be found in Edgar et al. (2001).

2.3 Dynamic Data Reconciliation

2.3.1 Linear Dynamic Data Reconciliation

In digital control, measurements are normally sampled at equal-spaced discrete time intervals. For a process that can be described by linear discrete models

$$\mathbf{x}_t = \mathbf{C}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_{t-1} + \mathbf{w} \tag{2.15}$$

$$\mathbf{y}_t = \mathbf{x}_t + \varepsilon \tag{2.16}$$

where $\mathbf{x}_t$ is a $S \times 1$ vector of state variables at time t, $\mathbf{y}_t$ is a $M \times 1$ vector of measurements at time t, $\mathbf{u}_{t-1}$ is a $R \times 1$ vector of manipulated inputs or disturbances to the process at time t-1, $\mathbf{w}$ is a $S \times 1$ vector of random variables that account for unknown disturbances corrupting the process, ε is a $M \times 1$ vector of random variables that represents the measurement errors, $\mathbf{B}$ and $\mathbf{C}$ are matrices of appropriate dimensions whose coefficients are known at all time. The optimal estimates for the true values of the state variables, $\mathbf{x}_t$, by the information provided by the measurements, $\mathbf{y}_t$, can be given by the *Kalman filter*. The Kalman filter was the earliest method used to recursively estimate state variables and process parameters for linear dynamic systems (Grewal et al., 1993). It has been widely applied in control for complex dynamic systems such as found in electronics manufacturing, aircraft, ships, and spacecraft. However, its practical application in chemical engineering has been relatively infrequent.

For the problem of reconciling process flow rates and storage volumes to calculate process inventories, Bagajewicz et al. (1997) recently presented an integral approach to solve this linear dynamic data reconciliation problem formulated as

$$\text{Minimize} \qquad J[\hat{\mathbf{y}}(t), \hat{\mathbf{z}}(t)] = [\mathbf{y}(t) - \hat{\mathbf{y}}(t)]^T \mathbf{V}^{-1} [\mathbf{y}(t) - \hat{\mathbf{y}}(t)] \tag{2.17}$$

subject to $\frac{d\hat{\mathbf{Q}}(t)}{dt} = \mathbf{A}\hat{\mathbf{s}}(t)$

$$\mathbf{C}\hat{\mathbf{s}}(t) = \mathbf{0}$$

where $\hat{\mathbf{Q}}(t)$ is a $R \times 1$ vector of reconciled values for continuous measurements of storage tank volumes, $\hat{\mathbf{s}}(t)$ is a $F \times 1$ vector of reconciled values for continuous measurements of flow rates, $\mathbf{y}(t)$ is the $M \times 1$ vector of the continuous measurements such that $\mathbf{y}(t) = [\mathbf{Q}(t), \mathbf{s}(t)]^T$ and $M = R + F$, $\mathbf{A}$ is a $R \times F$ incidence matrix corresponding to storage tanks as the nodes, and $\mathbf{C}$ is a $E \times F$ incidence matrix corresponding to splitters or mixers as the nodes.

The new matrices $\mathbf{D} = \begin{bmatrix} \mathbf{A} & -\mathbf{I} \\ \mathbf{C} & \mathbf{0} \end{bmatrix}$ and $\mathbf{S} = \begin{bmatrix} \hat{\mathbf{s}} & \frac{d\hat{\mathbf{Q}}}{dt} \end{bmatrix}^T$ were defined such that the model constraints can be written as $\mathbf{DS} = \mathbf{0}$. By partitioning matrix $\mathbf{D}$, redundant and nonredundant spatial variables were classified. The spatially nonredundant variables were eliminated by projection methods, then the spatially redundant variables were reconciled by an integral approach in which the dynamics of the spatially redundant variables were approximated by polynomials such that the model differential equations of the model constraints were integrated analytically. However, one difficulty of this approach is to determine the proper order of the polynomials.

2.3.2 Nonlinear Dynamic Data Reconciliation

In the case where nonlinear dynamic models are used as constraints on the measurements, Liebman et al. (1992) proposed an approach, called nonlinear dynamic data reconciliation (NDDR) to solve this problem. For discrete sampled measurements, the NDDR was formulated as

$$\text{Minimize} \qquad J(\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) = \sum_{t=0}^{c} [(\mathbf{y}_t - \hat{\mathbf{y}}_t)^T \mathbf{V}^{-1} (\mathbf{y}_t - \hat{\mathbf{y}}_t)] \tag{2.18}$$

$$\text{subject to} \qquad \mathbf{f}[\frac{d\hat{\mathbf{y}}(t)}{dt}, \hat{\mathbf{y}}(t), \hat{\mathbf{z}}(t)] = \mathbf{0}$$

$$\mathbf{g}[\hat{\mathbf{y}}(t), \hat{\mathbf{z}}(t)] = \mathbf{0}$$

$$\mathbf{h}[\hat{\mathbf{y}}(t), \hat{\mathbf{z}}(t)] \geq \mathbf{0}$$

where c is the time window width, **f** and **g** represent nonlinear differential algebraic equations (DAEs), and **h** represents the inequality constraints including simple upper and lower bounds. The NDDR method discretizes the nonlinear differential equations by the method of orthogonal collocation on finite elements, such that the differential equations are transformed into algebraic equations at each sampling time. After the discretization, the remaining problem is to minimize the least-squares objective function with the constraints of algebraic equalities and inqualities that can be efficiently solved by the nonlinear programming techniques at each sampling time. The moving window horizon approach to reduce the dimensions of the optimization problem was suggested by Liebman et al. (1992). Only measurements within the horizon (t = 0, 1, 2, ..., c) were reconciled during this NDDR run, and only the estimates for the current time

measurements, $\hat{\mathbf{y}}_c$, were used for online optimization and control purposes. The length of the moving window provides a means for tuning the performance of the NDDR algorithm. The bigger the magnitude of the window length, the better the performance of the NDDR algorithm is, but at the expense of larger computational time.

For the nonlinear dynamic data reconciliation problem (2.18), Ramamurthi et al. (1993) proposed a successively linearized horizon based estimator (SLHE). The DAEs in equation (2.18) were linearized about a nominal trajectory at each time step in the estimation horizon, and then, the linearized equations were discretized analytically as the constraints for the optimization problem. Analytical solutions for the estimates of the state variables at the beginning of the window horizon $\hat{\mathbf{y}}_{t=0}$ were obtained analytically with the linearized constraints. The estimates $\hat{\mathbf{y}}_t$ (t = 1, 2, ...c) within the window horizon were then obtained by integrating the nonlinear differential equations with the initial estimates $\hat{\mathbf{y}}_{t=0}$. A significant reduction in computational time was achieved using SLHE compared with NLP-based estimation strategy. One disadvantage of the SLHE approach is its inability to handle inequality constraints explicitly. Ramamurthi et al. (1993) reported that controller performance was improved by their control-relevant nonlinear dynamic data reconciliation method. However, quantitative assessment of the controller performance with the dynamic data reconciliation was not given in their work.

Albuquerque et al. (1996) applied implicit Runge-Kutta methods to discretize the DAEs in equation (2.18). The remaining problem was solved by nonlinear programming

techniques. They addressed several aspects of the dynamic data reconciliation problem that included gross-error detection and variable classification. For data reconciliation, they found that the uniqueness of the optimal solution and consequent convergence to the solution was closely tied to the observability and redundancy of the measurements that could be analyzed by the method of LU factorization.

Binder et al. (2002) discussed the dynamic data reconciliation (DDR) problem from the perspective of mathematical theory. They argued that the DDR problem is ill-posed by the constraints of the DAEs in (2.12), and therefore requires *regularization* of the DAEs. Actually, regularization has been implicitly applied either through discretization of continuous equations or by starting off with discrete dynamic models. They pointed out that appropriate discretization itself typically acts as regularization. The authors suggested a method that automatically generates sequences of discretization grids, based on upgrading approximate spaces until the best compromise between measurement data and regularization error was achieved. They demonstrated that non-equidistant grids of discretization offer advantages over uniform grids.

Several papers have reported applications of the NDDR approach in dynamic data reconciliation. Barbosa et al. (2000) applied the NDDR algorithm to a CSTR polymerization reactor. Soderstrom et al. (2000) applied a NDDR program code to a DCS to accurately track diluent inventories. Abu-el-zeet et al. (2002) applied NDDR tools including gross error detection to a model predictive control (MPC) scheme. The raw

measurements were processed by the NDDR module before being passed to the controllers. These strategies were implemented in a simulation of two continuous stirred tank reactors (CSTR) connected in series. It was shown that the overall performance of the controllers was improved significantly when the data were first reconciled prior to being fed to the controllers. In particular, the scheme effectively compensated for raw measurements contaminated by gross errors including systematic bias and outliers.

An alternative to solving the nonlinear dynamic data reconciliation problem is to use the extended Kalman filter (EKF). This extension typically involves replacing the nonlinear model equations with first-order approximations. Ahn et al. (1999) and Gao et al. (1999) recently tested the EKF approach in nonlinear model predictive control (NLMPC). The EKF-based NLMPC scheme was shown to have better performance than PID controllers. The disadvantage of the EKF is its inability to handle inequality constraints in the data reconciliation problem.

2.4 Summary

Steady-state data reconciliation is now a mature technology that has been successfully applied in various processing industries. However, research work will still continue and focus on the improvement of computational efficiency for the steady-state data reconciliation problem.

Dynamic data reconciliation technology is continuously evolving. For linear dynamic processes, the Kalman filter is recommended to give solutions to the estimation of the process state variables. For nonlinear dynamic processes without inequality constraints, the use of the extended Kalman filter should be considered. On the other hand, for nonlinear dynamic processes with inequality constraints, the NDDR approach should be used.

The Kalman filter should play an important role in the dynamic data reconciliation problems, because most processes can be constructed by linear dynamic models that are sufficient to capture the essence of process dynamics. In addition, empirical input-output linear dynamic models can be easily obtained via process identification either offline or online. On the other hand, nonlinear dynamic models are more difficult to obtain for chemical processes.

It should be noted that most previous work in data reconciliation assumed that the model constraints are perfect so that the reconciled data must satisfy perfect models. However, perfect models rarely exist and the reconciliation of measured values will more likely be a compromise between inaccuracies in both measurements and process models. Data reconciliation with uncertain models is a particular interest of this work.

NOMENCLATURE

$\mathbf{A}$: $C \times M$ incident matrix

$\mathbf{B}, \mathbf{C}$: Matrices in discrete dynamic models

$\mathbf{f}, \mathbf{g}, \mathbf{h}$: Vectors of model constraints

$\mathbf{P}$: $C \times C$ projection matrix

$\mathbf{s}(t)$: $F \times 1$ vector of continuous measurements of flow rates

$\mathbf{Q}(t)$: $R \times 1$ vector of continuous measurements of storage tank volumes

$\mathbf{u}$: $S \times 1$ vector of manipulated variables

$\mathbf{V}$: $M \times M$ covariance matrix of the measurements

$\mathbf{w}$: $S \times 1$ vector of process random disturbances

$\mathbf{x}$: $M \times 1$ vector of true values of process variables

$\mathbf{y}$: $M \times 1$ vector of raw measurements

$\mathbf{z}$: $N \times 1$ vector of estimates for unmeasured variables

Greek letters

ε: $M \times 1$ vector of random variables

λ: Penalty parameter

Superscripts and Subscripts

t: Time instant

$\wedge$: Estimate

References

Abu-el-zeet, Z.H.; P.D. Roberts; and V.M. Becerra, "Enhancing model predictive control by data reconciliation", AIChE, 48, 324-333, 2002.

Ahn, S.; M. Park; and H. Rhee, "Extended Kalman filter-based nonlinear model predictive control for a continuous MMA polymerization reactor", Ind. Eng. Chem. Res., 38, 3942-3949, 1999.

Albuquerque, J.; and L.T. Biegler, "Data reconciliation and gross error detection for dynamic systems", AIChE, 42, 2841-2856, 1996.

Bagajewicz, M.J.; and Q. Jiang, "Integral approach to plant linear dynamic reconciliation", AIChE, 43, 2546-2558, 1997.

Barbosa Jr, V.P.; M. R.M. Wolf; and R. Maciel Fo, "Development of data reconciliation for dynamic nonlinear system: application the polymerization reactor", Computers Chem. Engng, 24, 501-506, 2000.

Binder, T.; L. Blank; W. Dahman; and W. Marquardt, "On the regularization of dynamic data reconciliation problems", Journal of Process Control, 12, 557-567, 2002.

Edgar, T.F.; D.M. Himmelblau; and L. S. Lasdon, "Optimization of chemical processes", 2nd Ed., McGraw Hill, 2001.

Gao, F.; F. Wang; and M. Li, "A neural-network-based nonlinear controller using extended Kalman filter", Ind. Eng. Chem. Res., 38, 2345-2349, 1999.

Grewal, M.S.; and A.P. Andrews, "Kalman filtering, theory and practice", Prentice Hall, 1993.

Liebman, M. J.; T. F. Edgar; and L. S. Lasdon, "Efficient data reconciliation and estimation for dynamic process using nonlinear programming techniques", Computers Chem. Engng, 16, 963-986, 1992.

Narasimhan, S; and C. Jordache, "Data reconcilation & gross error detection, an intelligent use of process data", Gulf Publishing Company, Houston, 2000.

Ramamurthi, Y.; P.B. Sistu; and B.W. Bequette, "Control-relevant dynamic data reconciliation and parameter estimation", Computers Chem. Engng, 17, 41-59, 1993.

Romagnoli, J.A.; and M.C. Sanchez, "Data processing and reconciliation for chemical process operations", Academic Press, 2000.

Soderstrom, T.A.; T.F. Edgar; L.P. Russo; and R.E. Yong, "Industrial application of a large scale dynamic data reconciliation strategy", Ind. Eng. Chem. Res., 39, 1683-1693, 2000.

CHAPTER 3 (PAPER 1)

Closed-loop Data Reconciliation for the Control of a Binary Distillation Column

S. Bai, J. Thibault, and D. D. McLean

Department of Chemical Engineering
University of Ottawa, Ottawa, Canada K1N 6N5

Abstract

Data reconciliation is a procedure that makes use of process models along with process measurements to give more precise and consistent estimates for process variables. Data reconciliation has been traditionally used to provide a more representative set of data to calculate steady-state inventories and process yields. For dynamic systems, the use of data reconciliation is relatively nascent. This paper examines the potential use of data reconciliation in closed-loop control as a filter to attenuate the noise in measurements of the controlled variables so that the controllers can access more accurate sets of data. Data reconciliation filters were implemented in simulations of a PID control system for a binary distillation column. Results showed that data reconciliation could efficiently reduce the propagation of measurement noise in control loops, so that the overall performance of the controller is enhanced.

3.1 Introduction

Reliable knowledge of the current state of a process is paramount to its assessment, control and optimization. However, chemical plant measurements are inevitably corrupted by process noise. The impact of such noise on process measurements can be represented by

$$\mathbf{y} = \mathrm{E}[\mathbf{y}] + \varepsilon \tag{3.1}$$

where $\mathbf{y}$ is an $M \times 1$ vector of the M observed values of the measured variables, $E[\mathbf{y}]$ is the $M \times 1$ vector of expected values of these process variables and ε is the $M \times 1$ vector of the measurement noise for which it is usually assumed to be normally distributed with mean zero and known variance.

The measurement noise is usually of high frequency, and results in high-frequency oscillations of manipulated variables that deteriorate the performance of the control system. In order to approach optimal control, it is therefore necessary to attenuate the measurement noise before calculating the control actions. Analog and digital filters are widely used to cope with these problems. Classical digital filters, including the linear/nonlinear exponential filters, moving average filters, and polynomial filters, have been discussed by Narasimhan and Jordache (2000). These classical digital filters provide satisfactory performance under steady-state conditions if they are well tuned. However, under transient conditions, their efficiency to reduce noise is decreased and time delays are introduced. Today, the use of high-speed computers in process control makes it

possible to implement more complex algorithms to filter the measurement noise. One such approach is data reconciliation (DR).

Data reconciliation has been widely applied in a variety of processing industries. However, most applications have been limited to process monitoring, gross error detection and process inventory calculations. For example, Bussani et al. (1995), Pierucci et al. (1996), Chiari et al. (1997), and Li et al. (2001) respectively applied on-line steady-state data reconciliation and optimization to a hydrogen plant, olefin plant, refinery, and crude oil distillation unit. McBrayer et al. (1998) applied open-loop nonlinear dynamic data reconciliation to a feed-blend tank to estimate unmeasured flow rates. Soderstrom et al. (2000) successfully implemented a real-time dynamic data reconciliation strategy to improve inventory calculations of a diluent for a chemical plant.

Despite these applications of DR in the complex plants, no reports of investigations of the use of DR with widely used PID control schemes were found. The use of data reconciliation within the structure of control loops is schematically shown in Figure 3.1. The data reconciliation algorithm is treated as a digital filter to reconcile the raw measurements; then the reconciled data are used by the controllers to calculate values for the manipulated variables. Recently, Abu-el-zeet et al. (2002) investigated the use of dynamic data reconciliation and systematic bias detection in model predictive control (MPC).

The objective of this work was to evaluate the use of quasi-steady-state data reconciliation to improve the performance of conventional PID controllers when measurement noise is significant. The presence of measurement noise distorts the dynamics of the process, leading to the detuning of controllers that consequently results in more sluggish control action. In this work, data reconciliation algorithms, embedded within control loops, were used as filters to reduce the impact of measurement noise and its propagation through the control loop. It was expected that noise reduction would allow use of higher controller gains such that the performance of the controller would be enhanced. The performance of a quasi-steady-state data reconciliation filter for the control of a binary distillation column was evaluated.

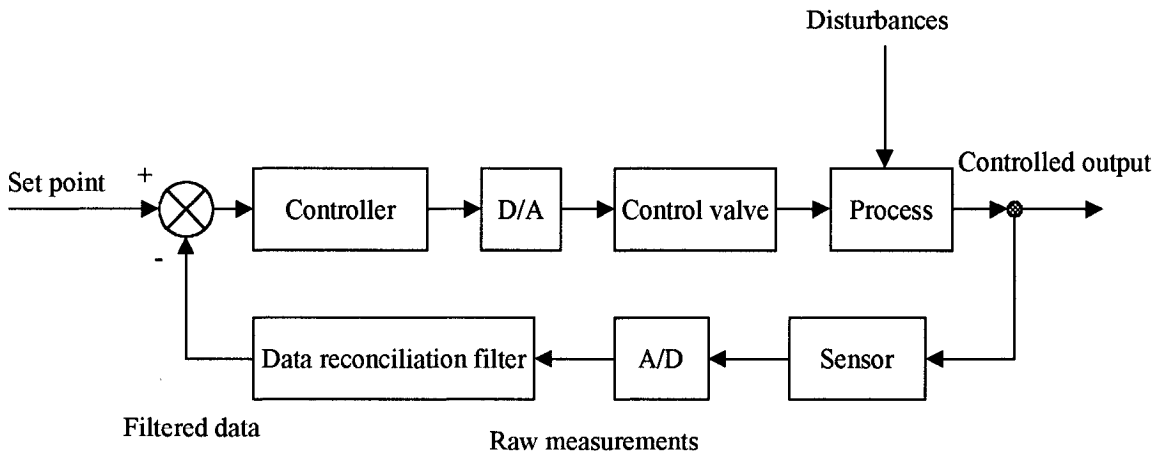

Figure 3.1 Application of data reconciliation algorithms in process control

3.2 Formulation of Data Reconciliation Filters

Data reconciliation techniques are used to obtain precise estimates of the true (expected) values of measured or unmeasured process variables that are consistent with underlying

process models. Data reconciliation can be formulated as a constrained weighted least-squares optimization problem such that reconciled values of process variables are those that

$$\begin{aligned} &\text{Minimize} && J(\hat{\mathbf{y}},\hat{\mathbf{z}}) = (\mathbf{y}-\hat{\mathbf{y}})^T \mathbf{V}^{-1}(\mathbf{y}-\hat{\mathbf{y}}) \\ &\text{subject to} && \mathbf{f}(\hat{\mathbf{y}},\hat{\mathbf{z}}) = \hat{\delta} \\ & && \mathbf{y}_l \le \hat{\mathbf{y}} \le \mathbf{y}_u \\ & && \mathbf{z}_l \le \hat{\mathbf{z}} \le \mathbf{z}_u \end{aligned} \tag{3.2}$$

where $\hat{\mathbf{y}}$ is the $M \times 1$ vector of estimates for the true values of the raw measurements, $\hat{\mathbf{z}}$ is the $N \times 1$ vector of estimates for the true values of the unmeasured variables, $\mathbf{V}$ is the $M \times M$ measurement covariance matrix. Vectors, $\mathbf{y}_l$, $\mathbf{y}_u$, $\mathbf{z}_l$, and $\mathbf{z}_u$ are the vectors of lower and upper bounds for the measured and unmeasured variables, $\mathbf{f}$ is the $C \times 1$ model constraint vector whose elements are process model equations (e.g., mass and energy balances) that may be algebraic or differential and $\hat{\delta}$ is the $C \times 1$ vector of model residuals that are the estimates of model random error $\mathbf{f}(\mathbf{y},\mathbf{z}) = \delta$.

The conventional data reconciliation algorithm formulated by equation (3.2) considers the process models as exact such that all elements of the model residual vector $\hat{\delta}$ are equal to zero. This implies that reconciled values are adjusted to perfectly satisfy these models. However, perfect models rarely exist and the reconciliation of measured values will more likely be a compromise between inaccuracies in both measurements and process models. In such cases, the reconciled values of the process variables at each

sampling time, $\hat{\mathbf{y}}_t$ and $\hat{\mathbf{z}}_t$, should be obtained by minimizing the weighted sum of squares of measurement and model errors defined as

$$\begin{aligned} \text{Minimize} \quad & J(\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) = (\mathbf{y}_t - \hat{\mathbf{y}}_t)^T \mathbf{V}^{-1} (\mathbf{y}_t - \hat{\mathbf{y}}_t) + \mathbf{f}^T (\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) \boldsymbol{\Omega}^{-1} \mathbf{f}(\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) \quad (3.3) \\ \text{subject to} \quad & \mathbf{y}_{t,l} \le \hat{\mathbf{y}}_t \le \mathbf{y}_{t,u} \\ & \mathbf{z}_{t,l} \le \hat{\mathbf{z}}_t \le \mathbf{z}_{t,u} \end{aligned}$$

where $\boldsymbol{\Omega}$ is the $C \times C$ covariance matrix of model error having the form

$$\boldsymbol{\Omega} = \begin{bmatrix} \nu_{1,1} & \nu_{1,2} & \cdots & \nu_{1,i} & \cdots & \nu_{1,C} \\ & \nu_{2,2} & \cdots & \nu_{2,i} & \cdots & \nu_{2,C} \\ & & \cdots & \cdots & \cdots & \cdots \\ & & & \nu_{i,i} & \cdots & \nu_{i,C} \\ & \textit{sym.} & & & \cdots & \cdots \\ & & & & & \nu_{C,C} \end{bmatrix}.$$

For linear models, the variances and covariances in the $\boldsymbol{\Omega}$ matrix can be obtained analytically. But for nonlinear models, they can be calculated by Monte Carlo simulation. If correlations between the model errors are neglected, $\boldsymbol{\Omega}$ becomes diagonal, and for nonlinear models the variances, $\nu_{i,i}$, can be approximated using first-order Taylor's expansions around the measurement values to give

$$\nu_{i,i} = \sum_{j=1}^{M} \left\{ \left(\left. \frac{\partial f_i}{\partial y_{t,j}} \right|_{\mathbf{y}_t} \right)^2 \sigma_j^2 \right\} \quad (3.4)$$

where σ_j^2 is the variance of the measurement y_j.

Equation (3.3) will be called the Single Set Data Reconciliation (SSDR) filter, since it only uses information available at the current sampling instant. It should be noted that: i) the process models, $\mathbf{f}(\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) = \hat{\delta}$, employed in the SSDR filter can be fundamental or empirical models; ii) since process models contain unmeasured variables, the number of models must be greater than the number of unmeasured variables in order that the models can provide the redundancy required for the data reconciliation in the SSDR case.

In order to further reduce the noise level, a moving average filter can be integrated within the structure of a SSDR filter to make use of past information about process measurements. If the differences of the current estimates with the most recent L measurements are minimized simultaneously with the minimization of the model errors, a Moving Window Data Reconciliation (MWDR) filter can be formulated as

$$\text{Minimize} \quad J(\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) = \sum_{k=0}^{L-1} \left[(\mathbf{y}_{t-k} - \hat{\mathbf{y}}_t)^T \mathbf{V}^{-1} (\mathbf{y}_{t-k} - \hat{\mathbf{y}}_t) + \mathbf{f}^T (\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) \Omega^{-1} \mathbf{f}(\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) \right] \quad (3.5)$$

$$\text{subject to} \quad \mathbf{y}_{t,l} \le \hat{\mathbf{y}}_t \le \mathbf{y}_{t,u}$$

$$\mathbf{z}_{t,l} \le \hat{\mathbf{z}}_t \le \mathbf{z}_{t,u}$$

where L is the window width in which L sets of recent measurements are used. The MWDR filter considers the process as being stationary within the window. The window width, L, can be considered a tuning parameter for the filter. If L is large, the filtered data should be smoother at the expense of longer time delays under dynamic changes.

When the variance of the model error is significantly larger than the variance of the measurement error, the MWDR filter approaches the simple moving average (MA) filter.

3.3 Simulation of Distillation Dynamics

3.3.1 Distillation System

The SSDR and MWDR filters were implemented in control of a binary (benzene/toluene) distillation column by simulation. The schematic diagram of the column is shown in Figure 3.2, and the geometric parameters of the column are summarized in Table 3.1.

Table 3.1 Column geometric parameters

Column diameter:	1.016 m	Weir height:	0.051 m
Tray type:	Sieve	Downcomer width:	0.132 m
Total number of trays:	21	Hole diameter:	4.8 mm
Tray spacing:	0.457 m	Hole pitch:	14 mm
Tray bubbling area:	0.68 m^2	Fractional hole area:	0.1
Downcomer area:	0.065 m^2	Reflux drum diameter:	1.016 m
Weir length:	0.681 m	Column base diameter:	1.016 m

The control configuration of the distillation column is as the followings. Two inferential feedback PI controllers, TIC-D and TIC-B, are used to control the top and bottom compositions of benzene by manipulating the reflux flow rate and the reboiler heat duty, respectively. Two other feedback PI controllers, LIC-D and LIC-B, are used to control the reflux drum and column-base liquid levels by manipulating the distillate flow rate and the bottom product flow rate, respectively. Three flow meters, FI-F, FI-D, and FI-B, monitor the feed, distillate, and bottom flow rates, respectively. The pressure meters, PI-

D and PI-B, measure the column top and bottom pressures. The nominal steady-state values for the measurements and their typical noise levels are listed in Table 3.2.

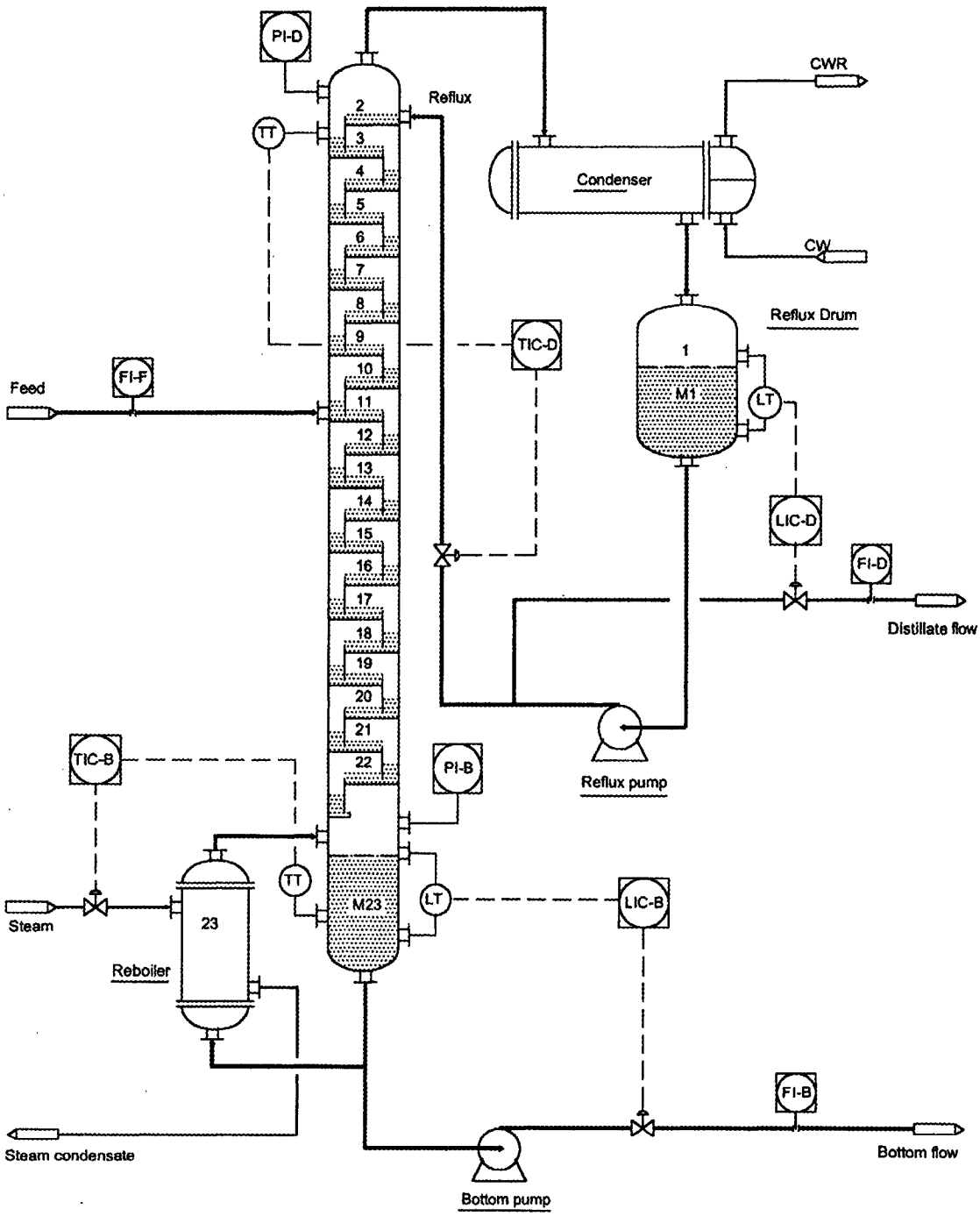

Figure 3.2 Schematic diagram of benzene-toluene distillation column

Table 3.2 Nominal steady-state values and noise levels of measured variables

Variable	Unit	Steady-state value	Standard deviation
Feed flow, (F)	kmol/h	100.0	0.50
Distillate flow, (D)	kmol/h	29.90	0.50
Bottom flow, (B)	kmol/h	70.10	0.50
To temperature, (T_D)	°C	84.20	0.25
Bottom temperature, (T_B)	°C	117.4	0.25
Reflux drum level, (H_D)	m	0.500	0.02
Column base level, (H_B)	m	0.700	0.02
Top pressure, (P_D)	kPa	111.2	0.50
Bottom pressure, (P_B)	kPa	125.6	0.50

Mathematical models for a distillation column are a set of differential-algebraic equations (DAEs) that describe the transient behavior of the column (Gani et al., 1986). These models include mass and heat balances, vapor-liquid phase equilibrium, tray hydraulics, physical properties and controller algorithms. The dynamic simulator of the column developed in this work was based on the following assumptions:

- The liquid holdups on trays, downcomers, reflux drum, and column base were considered to be well-mixed continuous stirred tanks.
- The vapor hold up was neglected.
- The tray efficiency was assumed to be 100%.
- The dynamics of the condenser, reboiler, measuring devices, and control valves were negligible compared with the dynamics of the distillation column.

The mass and heat balances for each stage (tray, condenser and reflux drum, reboiler and column base) are ordinary differential equations (ODEs), while the other phenomenological equations (e. g., phase equilibria and tray hydraulics) are algebraic

equations. The PI control algorithms are in difference form since the controllers are operated at discrete time. This set of mathematical models of the distillation column forms an ensemble of coupled DAEs that needs to be solved as a function of time. An explicit Euler's method was used to integrate the ODEs (Pantelides et al., 1992). The simulator was coded in FORTRAN.

3.3.2 Closed-loop Responses without Measurement Noise

For each pair of controlled and manipulated variables of a controller, the open-loop step response of the controlled variable to the manipulated variable, was first simulated to obtain the process reaction curve. The open-loop responses were approximated by either first-order-plus-dead-time, or pure-integrator-plus-dead-time models. The discrete sampling time, Δt, was set to 30 seconds for all measurements. Ziegler-Nichols tuning rules (Ogunnaike et al., 1994) were used to determine initial estimates of the discrete PI controller parameters, K_C and τ_I (see Table 3.3). The performance of the controllers was tested for disturbance regulation with the initial tuning parameters and without measurement noise. Results of this simulation are presented in Figure 3.3. These results illustrate that the controllers have good performance in the absence of measurement noise.

Table 3.3 Controller parameters

Controller	LIC-D		TIC-D		LIC-B		TIC-B	
	K_C kmol.h^{-1}.m^{-1}	τ_I s	K_C kmol.h^{-1}.°C^{-1}	τ_I s	K_C kmol.h^{-1}.m^{-1}	τ_I s	K_C MJ.h^{-1}.°C^{-1}	τ_I s
Initial Z-N values	-960.0	93.0	-87.0	96.6	-960.0	93.0	850.7	143.2
Detuned values	-35.8	131.3	-2.0	20.0	-59.8	40.9	44.9	19.3

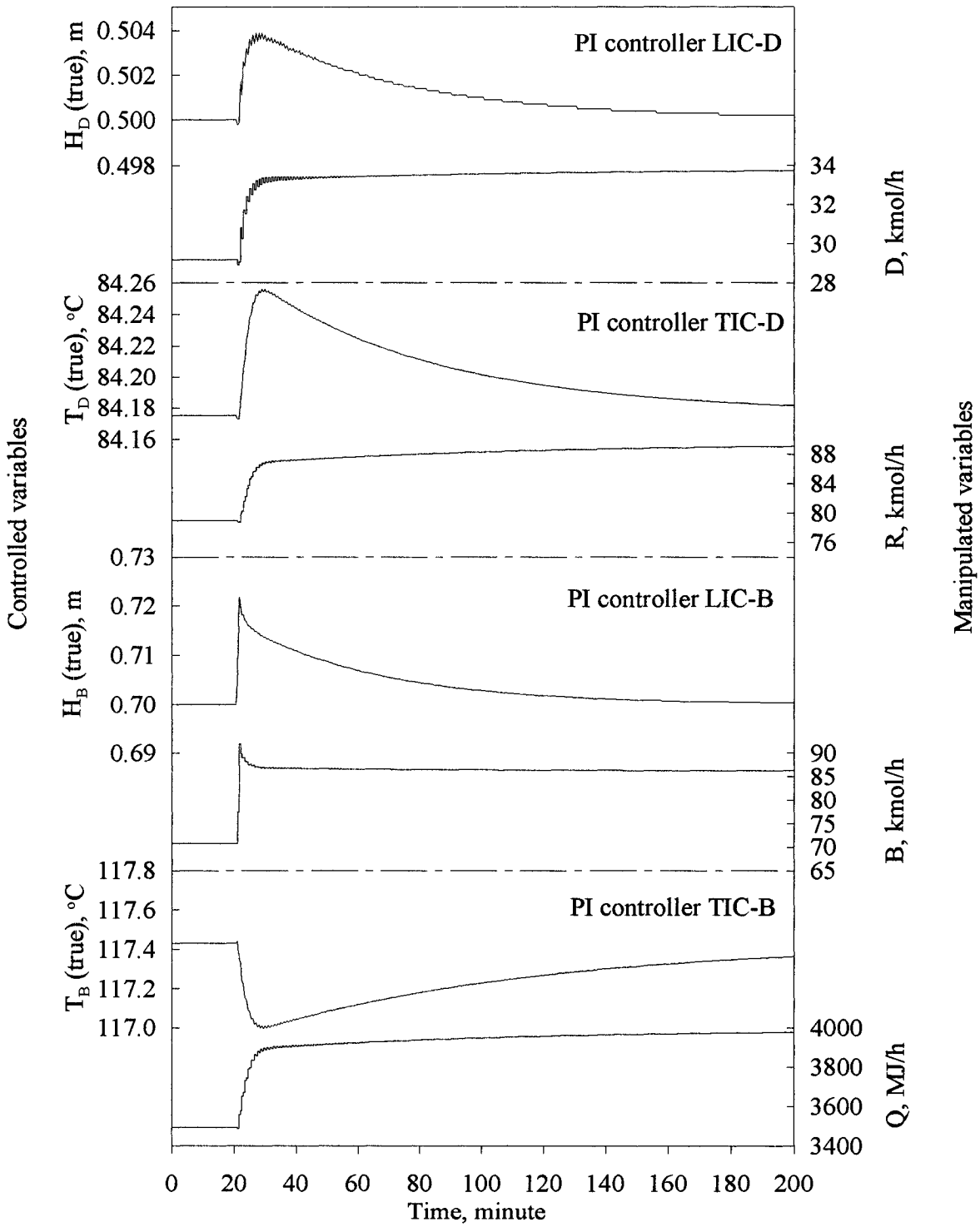

Figure 3.3 Closed-loop responses with initial Z-N tuning parameters without measurement noise for a 20% step increase in feed flow rate

3.3.3 Closed-loop Responses with Measurement Noise

In real plants all measurements are inevitably corrupted by noise. Using the standard deviations given in Table 3.2, Gaussian white noise was added to the true values of the process variables to provide measured values. Subsequently, these noisy measured values were used by the controllers to generate closed-loop responses. Results for the same feed flow disturbance are presented in Figure 3.4. The presence of the measurement noise has significantly reduced the performance of all control loops. The dynamic responses of the system are characterized by large, high-frequency, often saturated oscillations of manipulated variables that completely mask the expected responses following the feed flow rate disturbance.

To regain an acceptable system performance, controllers were retuned using the noisy measurements similar to those encountered in normal operation. For this multiple-input multiple-output (MIMO) discrete PI control system, the tuning objective function was formulated as

$$\text{Minimize } J(\mathbf{K_C}, \tau_\mathbf{I}) = \sum_{i=1}^{K} \left\{ W_i \sum_{t=1}^{t_s} (y^*_{t,i} - y_{t,i})^2 \Delta t + w_i \sum_{t=1}^{t_s} (u_{t,i} - u_{t-1,i})^2 \Delta t \right\} \qquad (3.6)$$

In equation (3.6), $J(\mathbf{K_C}, \tau_\mathbf{I})$ is defined as the cost function of the control system, $\mathbf{K_C}$ is the vector of controller gains, $\tau_\mathbf{I}$ is the vector of controller integral times, K is the total number of control loops, $y^*_{t,i}$ is the setpoint for controller i at time t, $u_{t,i}$ is the value of the manipulated variable for controller i at time t, W_i is the weighting factor for the integral of the squared error (ISE) of the controlled variable of controller i, w_i is the

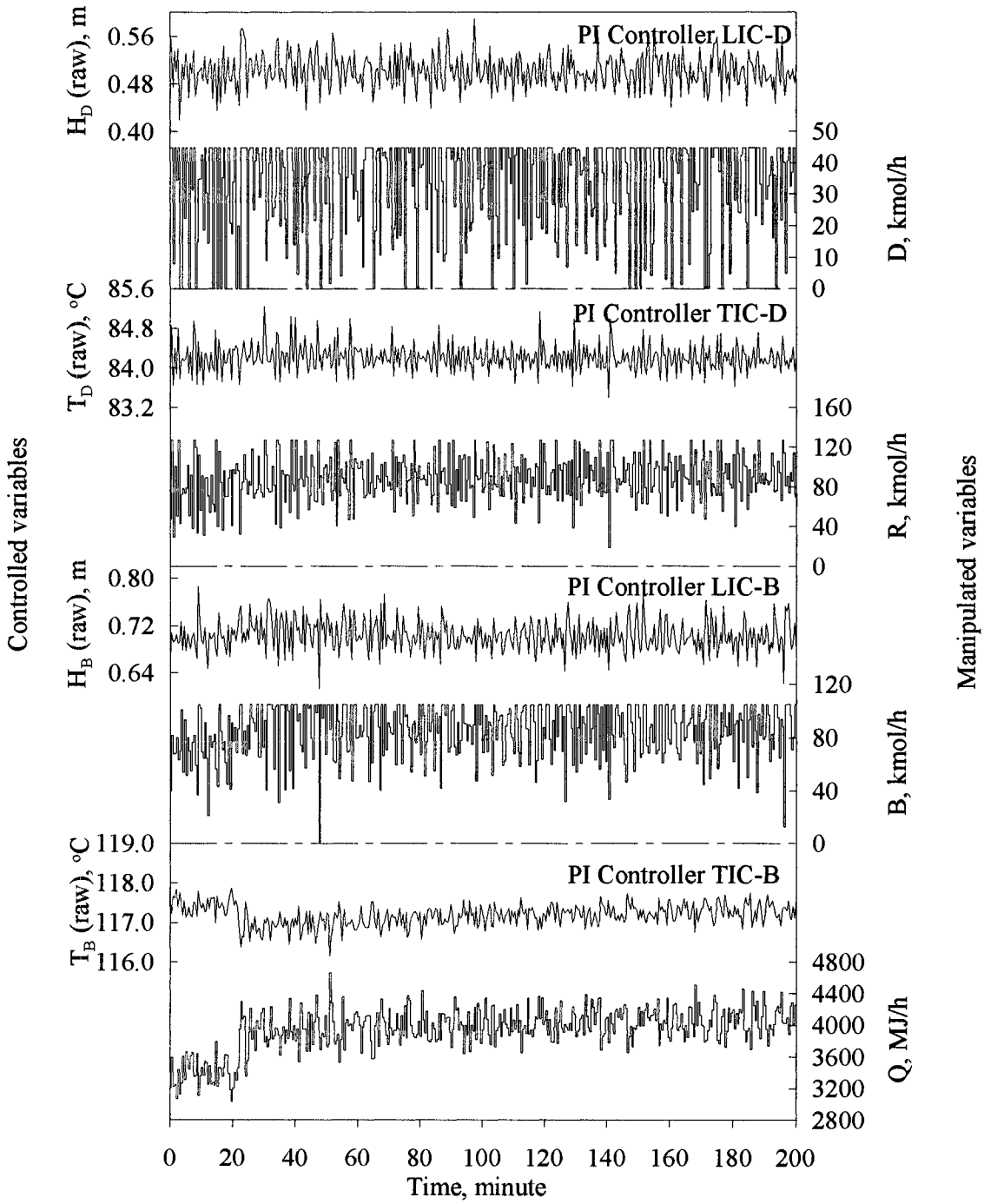

Figure 3.4 Closed-loop responses with initial Z-N tuning parameters with measurement noise levels listed in Table 3.2, for a 20% step increase in feed flow rate

weighting factor for the integral of the squared differences ($ISDU$) of the values of manipulated variable between times t and $t-1$ for controller i, and t_s represents a time sufficiently long to return to the nominal steady state.

In order to bring the scales of the *ISE* and *ISDU* components in (3.6) to similar levels, the *ISE* and *ISDU* for each controller were scaled by the initial values (ISE_0 and $ISDU_0$) obtained using the initial Z-N tuned controllers. The values of the weights W_i and w_i for were set to 0.1 and 0.1 for controller LIC-D, 0.2 and 0.1 for controller TIC-D, 0.05 and 0.05 for controller LIC-B, and 0.3 and 0.1 for controller TIC-B. The optimization was performed in two steps: a rough grid search was carried out to identify the optimal region, followed by a rigorous optimization using a quasi-Newton method.

Values of new optimal controller parameters are presented in Table 3.3 along with the original Z-N values. The ratio of the new K_C values to the original Z-N values ranges from 0.02 to 0.06 meanwhile the ratio of the τ_I values ranges from 0.14 to 1.41. The controller gains are on the average 22 times smaller whereas the integral times are on the average 1.8 times smaller. Closed-loop responses with the new controller parameters are presented in Figure 3.5. The oscillations of the manipulated variables have been significantly reduced. The dynamics of the controlled variables can be observed despite the noisy measurements, but at the expense of larger deviations from their setpoint during the transient period. It takes about 160 minutes for the column to return its steady state following the disturbance. The controllers become sluggish with the dutuned parameters.

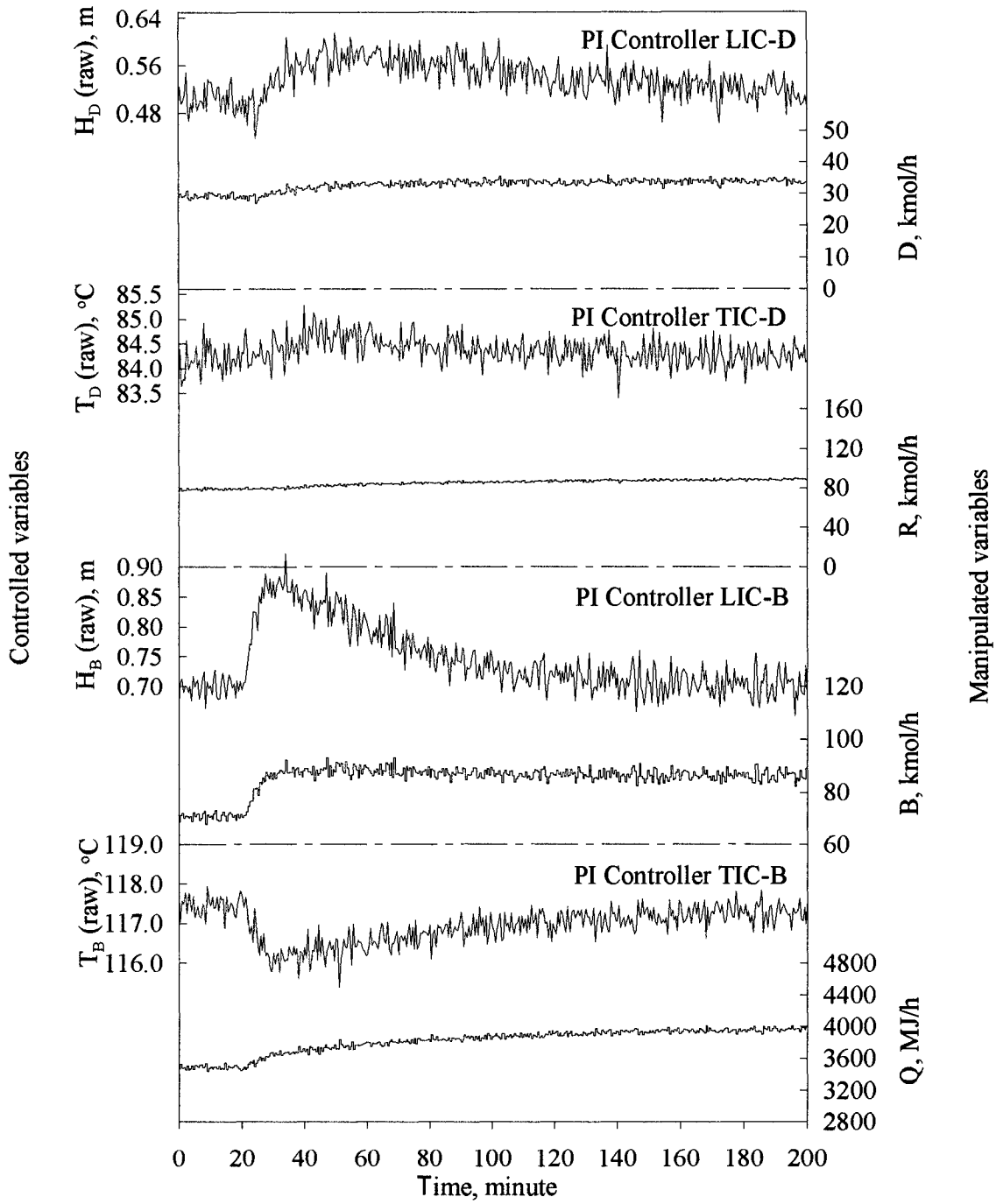

Figure 3.5 Closed-loop responses with detuned controllers with measurement noise level listed in Table 3.2, for a 20% step increase in feed flow rate

The effects of the variations of the manipulated variables on the true values of process variables were investigated. Figure 3.6 shows the true values of the four controlled variables at process nominal steady state when the four control loops were closed with noisy measurements but without external disturbances. The variances of the true values of controlled variables are significantly larger when the manipulated variables had larger oscillations with the initial Z-N tuning parameters. On the other hand, the variances of the true values of controlled variables are relatively small when the manipulated variables have small variations using the detuned controllers. The calculated variances of the controlled variables with the initial Z-N tuning parameters are on average 40 times greater than those of the detuned controllers. It is nevertheless important to attenuate unnecessary variations of the manipulated variables caused by process noise.

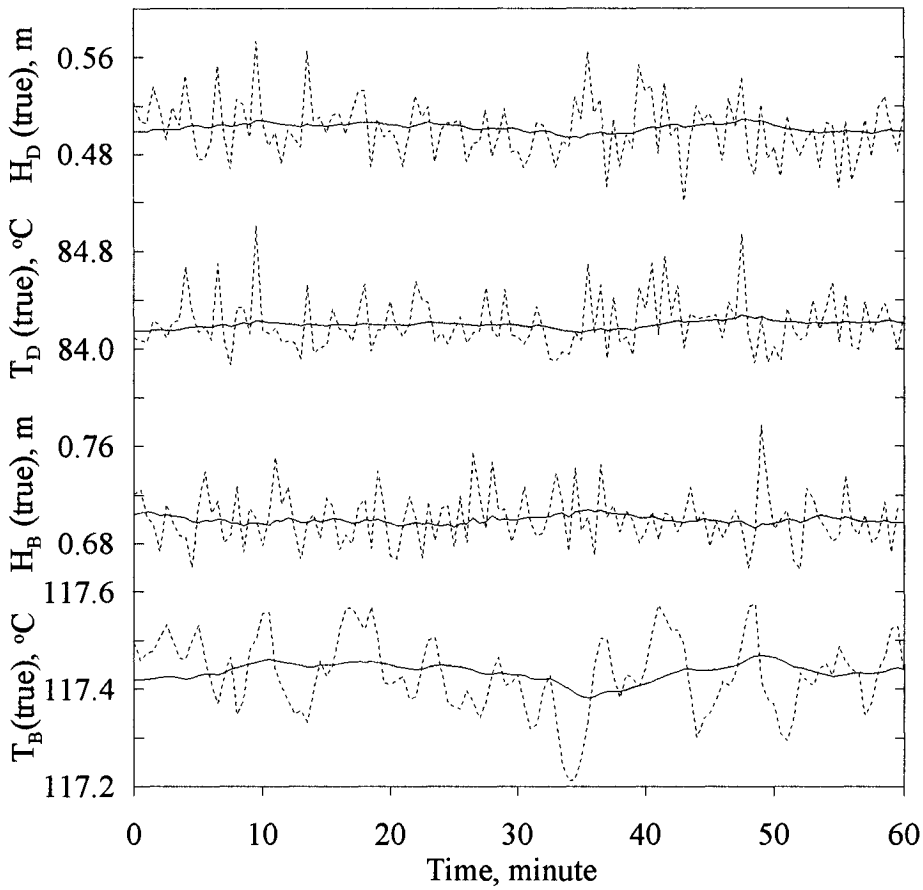

Figure 3.6: True values of controlled variables at process nominal steady-state conditions

..... With initial Z-N tuning controllers

——— With detuned controllers

3.4 Implementation of Data Reconciliation filters

To attenuate the propagation of the measurement noise, the SSDR and the MWDR filters formulated by equations (3.3) and (3.5) were implemented in the feedback control loops of the distillation column. The filtered data, at each instant t, were obtained by minimizing the objective function of (3.3) and (3.5), respectively. For the distillation column, the measurement vector is $\mathbf{y}_t = [F_t, D_t, B_t, T_{D,t}, T_{B,t}, H_{D,t}, H_{B,t}, P_{D,t}, P_{B,t}]^T$, and the vector of model residuals is given by

$$\mathbf{f}[\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t] = \begin{bmatrix} \hat{F}_t - \hat{D}_t - \hat{B}_t - \dfrac{(\hat{H}_{D,t} - \hat{H}_{D,t-1})A_D\rho_D}{\Delta t} - \dfrac{(\hat{H}_{B,t} - \hat{H}_{B,t-1})A_B\rho_B}{\Delta t} \\ \hat{F}_t x_F - \hat{D}_t \hat{x}_{D,t} - \hat{B}_t \hat{x}_{B,t} - \dfrac{(\hat{H}_{D,t}\hat{x}_{D,t} - \hat{H}_{D,t-1}\hat{x}_{D,t-1})A_D\rho_D}{\Delta t} - \dfrac{(\hat{H}_{B,t}\hat{x}_{B,t} - \hat{H}_{B,t-1}\hat{x}_{B,t-1})A_B\rho_B}{\Delta t} \\ \dfrac{\hat{P}_{D,t}\hat{x}_{D,t}}{10^{\left(A_1 - \frac{B_1}{\hat{T}_{D,t}+C_1}\right)}} + \dfrac{\hat{P}_{D,t}(1-\hat{x}_{D,t})}{10^{\left(A_2 - \frac{B_2}{\hat{T}_{D,t}+C_2}\right)}} - 1 \\ \dfrac{10^{\left(A_1 - \frac{B_1}{\hat{T}_{B,t}+C_1}\right)}\hat{x}_{B,t}}{\hat{P}_{B,t}} + \dfrac{10^{\left(A_2 - \frac{B_2}{\hat{T}_{B,t}+C_2}\right)}(1-\hat{x}_{B,t})}{\hat{P}_{B,t}} - 1 \end{bmatrix} \tag{3.7}$$

The first two rows of the model residuals are the equations of conservation of total mass and of one of the components, respectively. The assumption of process stationarity was partly relaxed by adding accumulation terms for both the reflux drum and the reboiler. On the other hand, the mass and component accumulations on each tray were neglected. A_D and A_B are the cross-sectional areas of the reflux drum and the column base. ρ_D and

ρ_B are the top and bottom liquid densities, which are assumed to be constant, and x_F is the feed composition. The third row is the dew point equation of the vapor leaving tray 2 employed to estimate the unmeasured variable $x_{D,t}$ using the measurements $P_{D,t}$ and $T_{D,t}$. The fourth row is the bubble point equation for the bottom liquid, used to estimate the unmeasured variable $x_{B,t}$ using measurements $P_{B,t}$ and $T_{B,t}$. Parameters A_j, B_j, and C_j are Antoine constants.

The random error of the four models was assumed uncorrelated, and the variances of the individual model errors were calculated by Monte Carlo simulation. The minimization of the objective function for the data reconciliation filter was implemented using a quasi-Newton method. The filtered data vector for the raw measurements obtained by optimization is $\hat{\mathbf{y}}_t = [\hat{F}_t, \hat{D}_t, \hat{B}_t, \hat{T}_{D,t}, \hat{T}_{B,t}, \hat{H}_{D,t}, \hat{H}_{B,t}, \hat{P}_{D,t}, \hat{P}_{B,t}]^T$, where the reconciled values of the controlled variables, $\hat{H}_{D,t}$, $\hat{T}_{D,t}$ $\hat{H}_{B,t}$, and $\hat{T}_{B,t}$, were used for the four controllers to calculate the manipulated variables. The auxiliary vector of the estimates of unmeasured variables is $\hat{\mathbf{z}}_t = [x_{D,t}, x_{B,t}]^T$.

The performance of the SSDR filter and the MWDR filter having a window width of 4 was evaluated for a 20% step increase of the feed flow rate. Results are presented in Figures 3.7 and 3.8, respectively. The SSDR filter is effective in reducing the noise propagation for the two liquid level control loops, whereas its filtering effects on the two temperature control loops are hardly noticeable. These results indicate that the mass balances play a predominant role in the data reconciliation filter with a secondary role for

thermodynamic models. On the other hand, the MWDR filter is effective for all control loops, and noise reductions through the control loops are significant. The overall performance of MWDR filter is better than that of the SSDR filter.

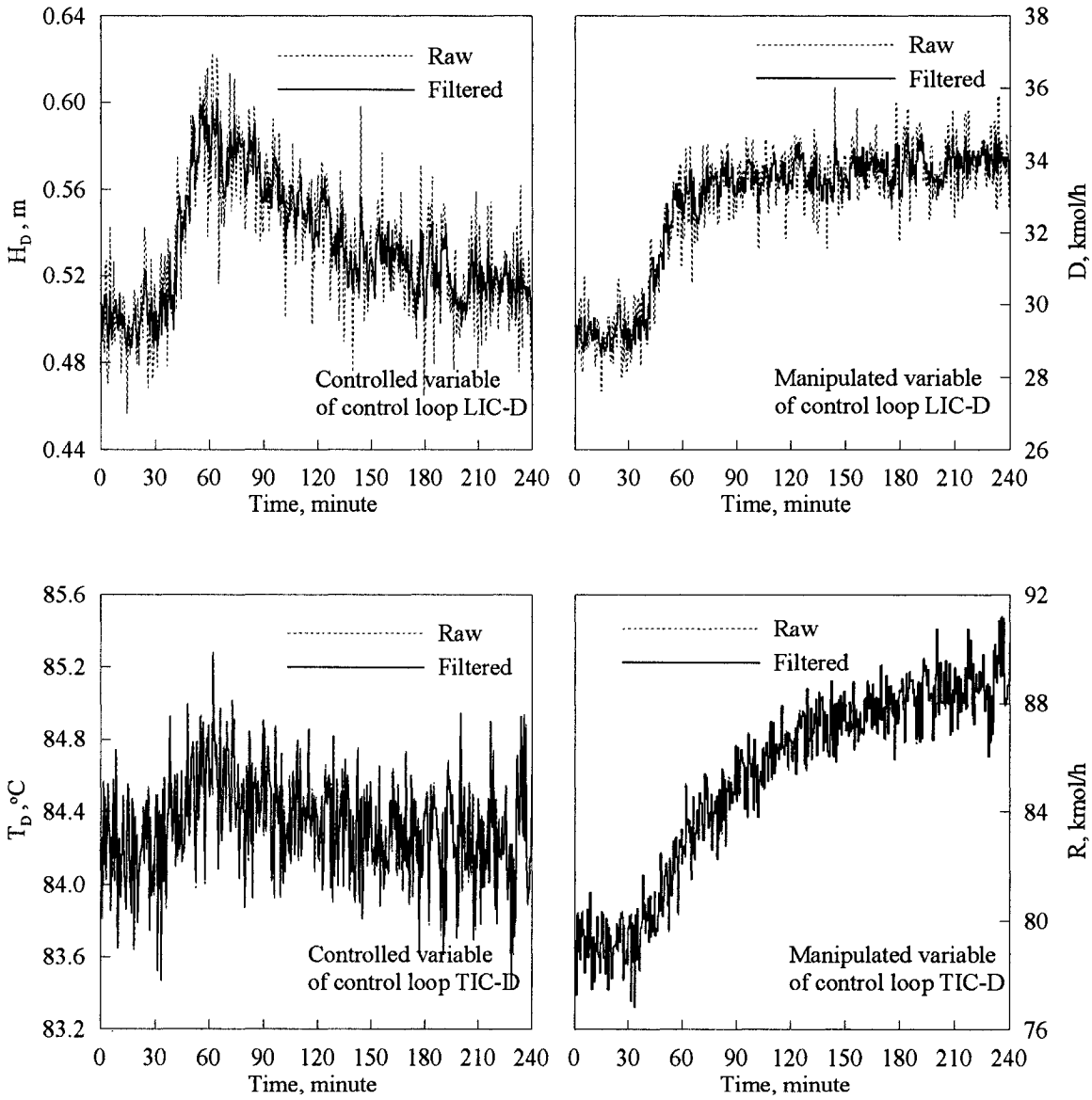

Figure 3.7a: Effectiveness of SSDR filter on control loops for 20% step increase in feed flow rate

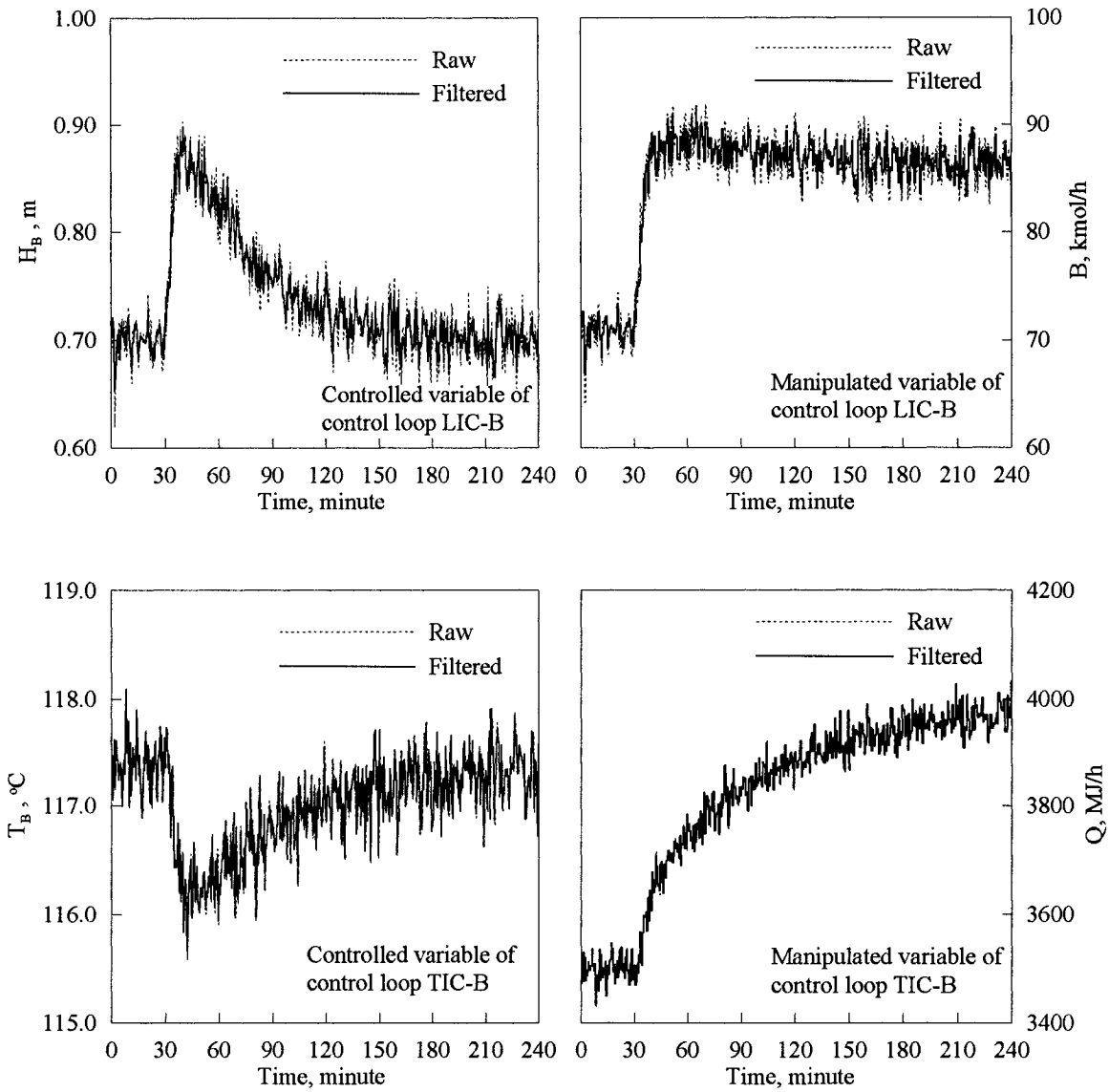

Figure 3.7b: Effectiveness of SSDR filter on control loops for 20% step increase in feed flow rate

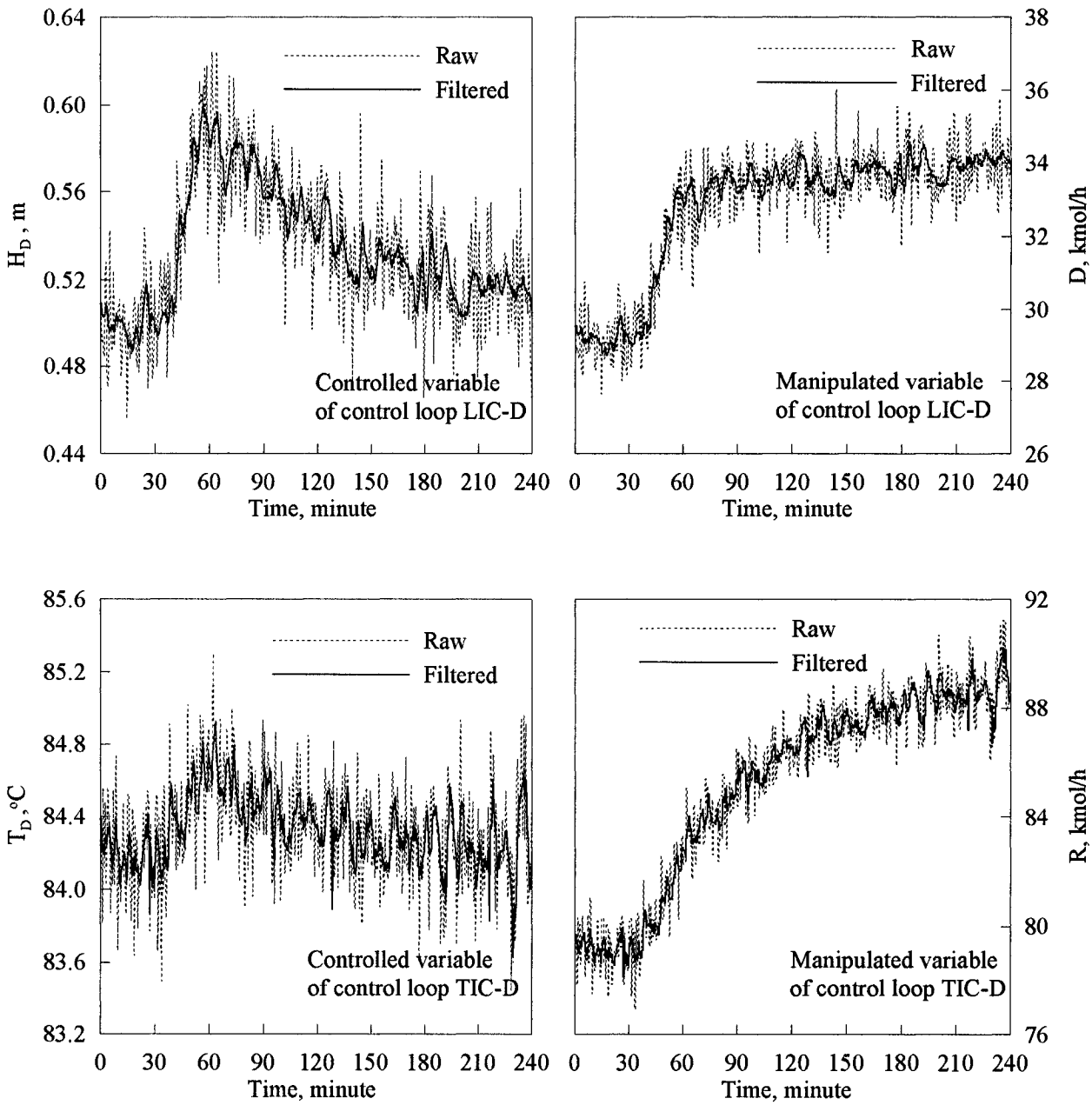

Figure 3.8a Effectiveness of MWDR filter on control loops for 20% step increase in feed flow rate

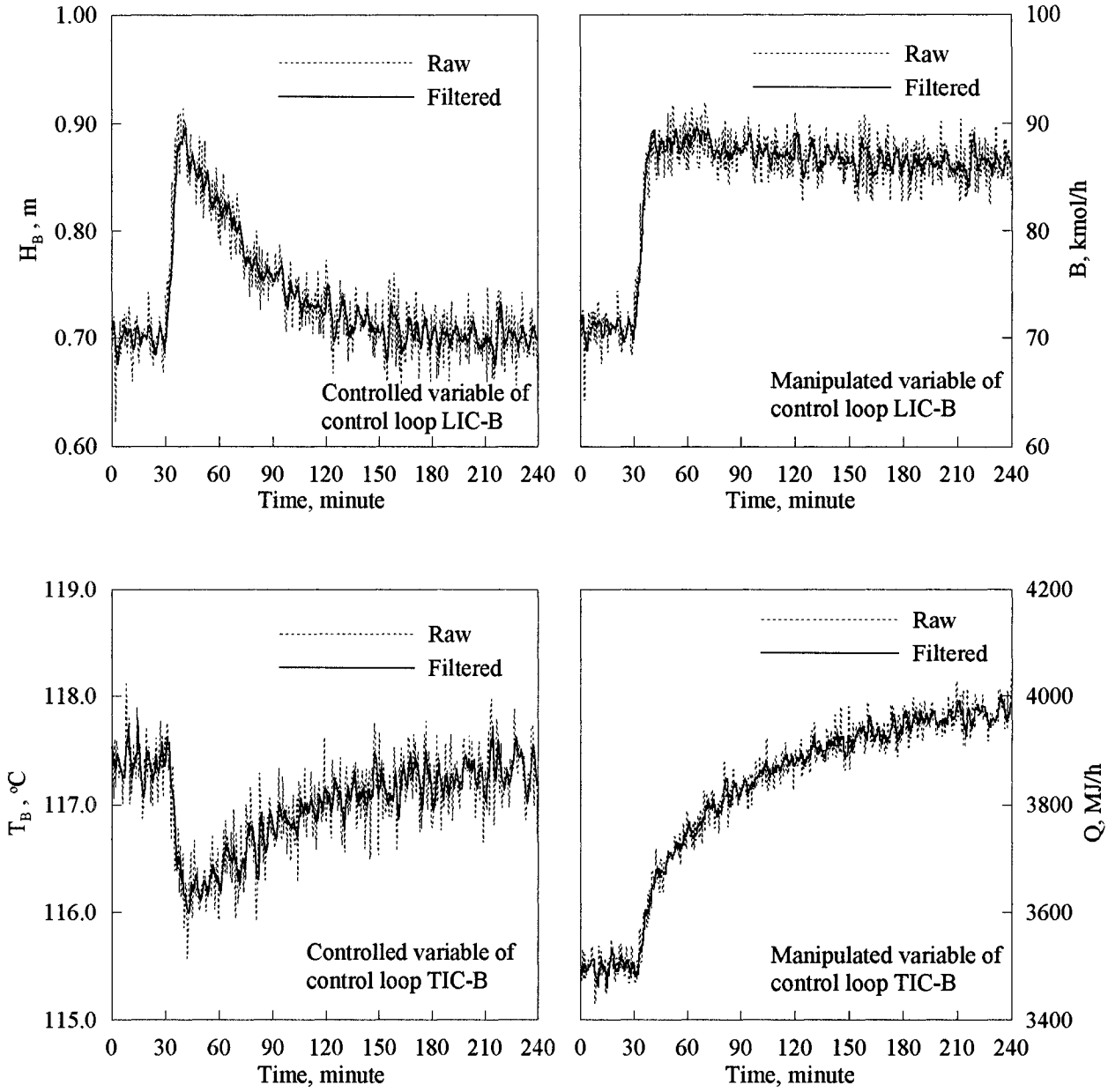

Figure 3.8b Effectiveness of MWDR filter on control loops for 20% step increase in feed flow rate

Values of *ISE* and *ISDU* calculated for each control loop, along with the evaluations of the cost function for each controller without and with DR filters are presented in Cases A, B and C of Table 3.4, respectively. The values of *ISDU* for individual control loops were reduced significantly with approximate one order of magnitude with the exception

of the two temperature control loops with the SSDR filter. On the other hand, the values of *ISE* remained relatively constant for the two filters. The cost function of the controller decreases due to the reduction in the *ISDU*. The overall cost function of the control system was reduced from 0.736 to 0.651 when the SSDR filter was used, and to 0.554 with the MWDR filter.

Due to the presence of measurement noise, it was necessary to considerably detune all control loops to obtain acceptable performance. Since data reconciliation filters were able to reduce the effect of noise, more aggressive controllers could therefore be used. To evaluate the performance of the control loops using a MWDR filter and with larger controller gains, the process was subjected to a 20% step increase of the feed flow rate. Results are presented in Figure 3.9 where the values of *ISE* and *ISDU* of each control loop are plotted as a function of the controller gain ratio K_C/K_{C0}. K_{C0} corresponds to the values of the optimal gains listed in Table 3.3 when process noise was considered. The gains of all control loops were incremented simultaneously. Figure 3.9 illustrates that the *ISE* decreases significantly when the control gains are increased by a factor of 4, and remains relatively constant for controller gain ratios in the range of 4 to 8. On the other hand, the values of *ISDU* increase exponentially with the increase of gain ratio as shown in Figure 3.9.

Table 3.4 Performance of controllers with data reconciliation filters ($J_i = W_i ISE_i + w_i ISDU_i$)

Case	DR Filter	LIC-D (W=0.1, w=0.1)			TIC-D (W=0.2, w=0.1)			LIC-B (W=0.05, w=0.05)			TIC-B (W=0.3, w=0.1)			$\sum_{k=1}^{4} J_k$
		ISE_1	$ISDU_1$	J_1	ISE_2	$ISDU_2$	J_2	ISE_3	$ISDU_3$	J_3	ISE_4	$ISDU_4$	J_4	
A	No	0.925	0.623	0.155	0.612	0.368	0.159	0.408	1.201	0.080	0.979	0.489	0.342	0.736
B	SSDR	0.929	0.098	0.103	0.618	0.373	0.161	0.420	0.489	0.045	0.984	0.471	0.342	0.651
C	MWDR	0.941	0.009	0.095	0.640	0.034	0.131	0.443	0.047	0.025	0.997	0.033	0.303	0.554
D	MWDR	0.144	0.128	0.027	0.186	0.351	0.072	0.052	0.652	0.035	0.222	0.402	0.107	0.241

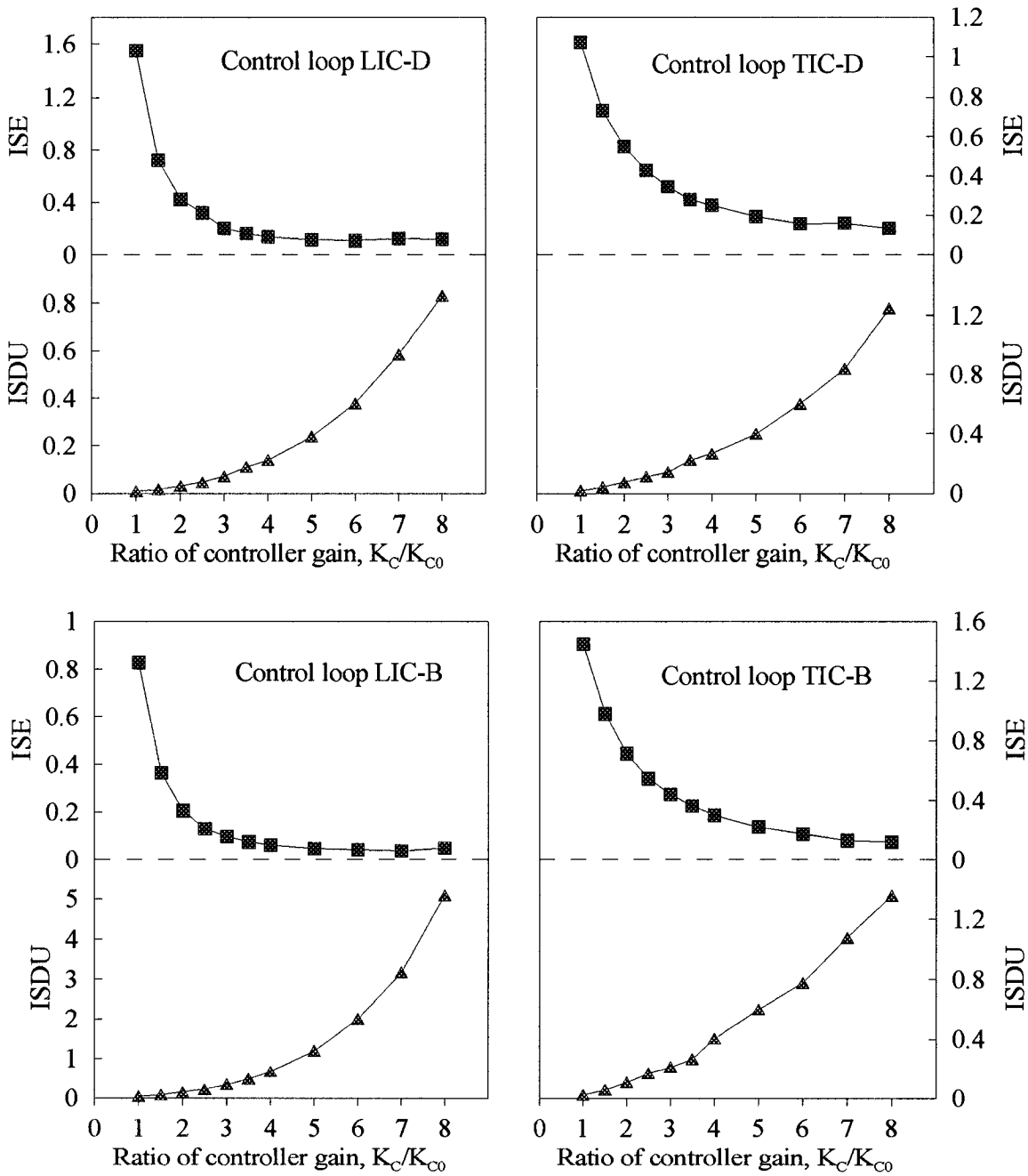

Figure 3.9: Performance of control loops with MWDR filter with higher controller gains (K_{C0}: detuned controller gains listed in Table 3.3)

The cost function for each control loop with the increase of K_C/K_{C0} was calculated, and the results are presented in Figure 3.10. For the control loops LIC-D, TIC-D, LIC-B and TIC-B, their minimum values of cost function were obtained at $K_C/K_{C0} = 3.0$, $K_C/K_{C0} = 4.0$, $K_C/K_{C0} = 2.5$ and $K_C/K_{C0} = 5.0$, respectively. However, the minimum value of the overall cost function of the control system was calculated at $K_C/K_{C0} = 3.5$ for all controllers. Using the new optimal controller gains, the values of *ISE*, *ISDU*, and cost function for each control loop were evaluated and the results are listed in Case D of Table 3.4. The overall cost function of the control system was reduced from 0.554 to 0.241 when the new optimal controller gains were used with the MWDR filter.

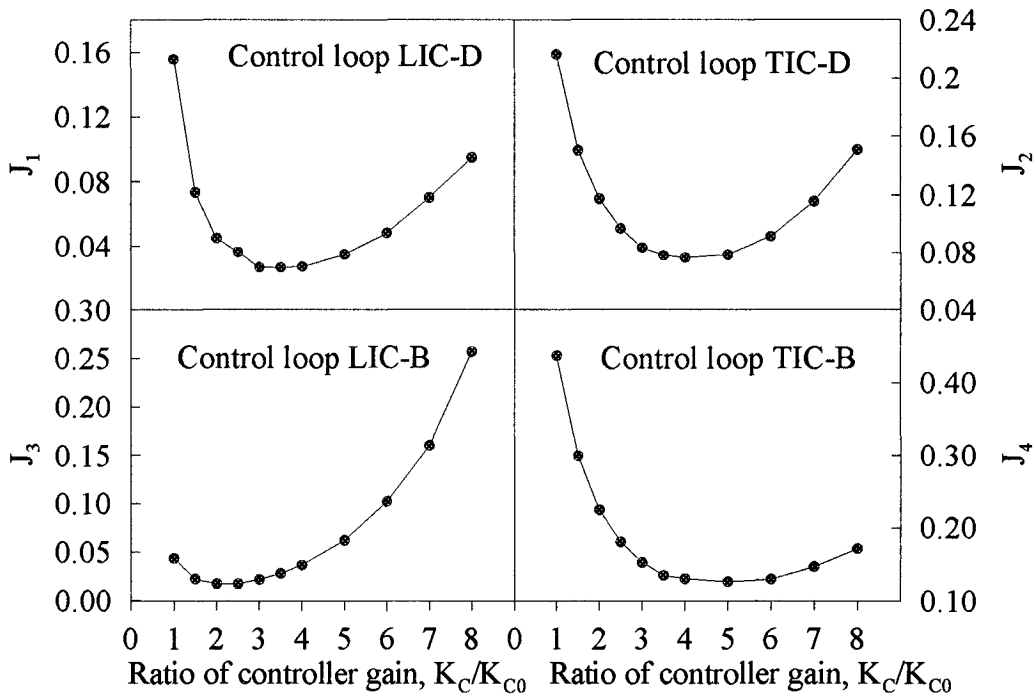

Figure 3.10: Cost function of control loop with MWDR filter with higher controller gains (K_{C0}: detuned controller gains listed in Table 3.3)

3.5 Conclusion

The magnitude of measurement noise must be considered when tuning controllers using models of the process. Otherwise, the controllers are too aggressive resulting in excessive variations in the controlled and manipulated variables. On the other hand, the detuning of controllers due to the presence of noise leads to more sluggish responses and lower controller performance. Controller performance can be partly regained using data reconciliation filter in the feedback loop to obtain more consistent values of the controlled variables.

In this investigation, two data reconciliation filters were used: the SSDR filter using only the current set of measurements, and the MWDR filter that makes use of a series of past measurements to further decrease the effect of noise. It was shown that DR filters can significantly attenuate the effect of noise which allowed to use more aggressive controllers and thereby improve the overall performance of all control loops.

The quasi-steady-state data reconciliation filter developed in this work employed measurement spatial redundancy for which more process space variables need to be measured to provide redundant information. Future work will focus on dynamic data reconciliation in which input-output transfer function models will be developed for the closed-loop data reconciliation algorithm.

Nomenclature

A:	Cross-sectional area, m^2
A_j, B_j, C_j:	Antoine constants
B:	Distillation bottom flow rate, kmol/h
C:	Number of process models
D:	Distillate flow rate, kmol/h
$E[\mathbf{y}]$:	$M \times 1$ vector of expected values of process variables
$\mathbf{f}$:	$C \times 1$ model constraint vector
F:	Feed flow rate, kmol/h
H:	Liquid level, m
$\mathbf{K_C}$:	Vector of the controller gain
K:	Total number of control loops
L:	Window width
M:	Number of measured variables
N:	Number of unmeasured variables
P:	Pressure, kPa
Q:	Reboiler heat duty, MJ/h
R:	Reflux flow rate, kmol/h
Δt:	Discrete sampling time interval, s
T:	Temperature, °C
u:	Manipulated variable
$\mathbf{V}$:	$M \times M$ covariance matrix of the measurements
W:	Weighing factor for ISE
w:	Weighting factor for ISDU
x:	Composition, mole fraction
$\mathbf{y}$:	$M \times 1$ vector of raw measurements
$\mathbf{z}$:	$N \times 1$ vector of estimates for unmeasured variables

Greek letters

ε:	$M \times 1$ vector of random variables
σ^2:	Variance
δ:	$C \times 1$ vector of model random error
$\mathbf{\Omega}$:	$C \times C$ covariance matrix of model error
τ_{I}:	Vector of the controller integral time
ρ:	Liquid density, kg/m^3

Superscripts and Subscripts

B:	Column bottom
D:	Distillate
F:	Feed
l:	Lower bound
t:	Time instant
u:	Upper bound
$\wedge$:	Estimate
*:	Setpoint

Acronyms

DAEs:	Differential-algebraic equations
ISE:	Integral of squared errors for controlled variable
ISDU:	Integral of squared difference of manipulated variable
MIMO:	Multiple-input multiple-output
MA:	Moving average
MWDR:	Moving window data reconciliation
MPC:	Model predictive control
ODEs:	Ordinary differential equations
SSDR:	Single set data reconciliation

References

Abu-el-zeet, Z.H.; P.D. Roberts; and V.M. Becerra, "Enhancing model predictive control by data reconciliation", AIChE J., 48, 324-333, 2002.

Bussani, G.; M. Chiari; M.G. Grottoli; S. Pierucci; T. Faravelli; G. Ricci; and G. Gioventu, "Application of data reconciliation and optimization procedure to hydrogen plant", Computers Chem. Engng, 19, Suppl. S299-S304, 1995.

Chiari, M.; G. Bussani; M.G. Grottoli; and S. Pierucci, "On-line data reconciliation and optimization: Refinery applications", Computers Chem. Engng, 21, Suppl. S1185-S1190, 1997.

Gani, R.; C. A. Ruiz; and I.T. Cameron, "A generalized model for distillation columns – I: model description and applications", Computers Chem. Engng, 10, 81-198, 1986.

Li, B.; B. Chen; J. Jiang; and S. Cong, "Steady-state online data reconciliation in a crude oil distillation unit", Hydrocarbon Processing, March 2001, 61-64.

McBrayer, K.F.; T.A. Soderstrom; T.F. Edgar; and R.E. Young, "The application of nonlinear dynamic data reconciliation to plant data", Computers Chem. Engng, 22, 1907-1911, 1998.

Narasimhan, S; and C. Jordache, "Data reconcilation & gross error detection, an intelligent use of process data", Gulf Publishing Company, Houston, 2000.

Ogunnaike, B.A.; and W.H. Ray, "Process dynamics, modeling, and control", Oxford University Press, 1994.

Pantelides, C.C.; and P.I. Barton, "Equation-oriented dynamic simulation: current status and future perspectives", Supplement to Computers & Chemical Engineering, European Symposium on Computer Aided Process Engineering-2, October 1992.

Pierucci, S.; P. Brandani; E. Ranzi; and A. Sogaro, "An industrial application of an on-line data reconciliation and optimization problems", Computers Chem. Engng, 20, Suppl., S1539-S1544, 1996.

Soderstrom, T.A.; T.F. Edgar; L.P. Russo; and R.E. Young, "Industrial application of a large-scale dynamic data reconciliation strategy", Ind. Eng. Chem. Res. 39, 1683-1693, 2000.

CHAPTER 4 (PAPER 2)

Enhancing Controller Performance via Dynamic Data Reconciliation

S. Bai, D. D. McLean, and J. Thibault

Department of Chemical Engineering

University of Ottawa, Ottawa, Canada K1N 6N5

Abstract

Chemical process measurements contain some degree of error that can be random or systematic. The presence of measurement noise usually results in detuned controllers in order to prevent excessive high-frequency variations of manipulated variables. However, with detuned controllers, the process is characterized by sluggish dynamic responses. In this paper, a Dynamic Data Reconciliation (DDR) algorithm, embedded within the structure of feedback control loops, is used to reconcile noisy measurements. The DDR algorithm employs discrete dynamic models, both phenomenological and empirical, as constraints in reconciling noisy measurements. Models are not assumed to be exact such that the data reconciliation algorithm is based on a compromise between the actual measurements and the model predicted values. Simulation results for a storage tank and a distillation column under PI control demonstrate that using a DDR algorithm within the structure of a feedback control loop not only results in faster responses of controlled variables, but also reduces variations of manipulated variables. Consequently, the overall performance of the control system is enhanced.

4.1 Introduction

Chemical process measurements contain some degree of error that can be random or systematic. When erroneous measurements are used in process monitoring, control, scheduling and optimization, the state of the process is more or less misrepresented. Reliable information about the state of the process is necessary, especially with the implementation of distributed control systems in which a large number of on-line measurements are taken and used for process control and optimization purposes. Data reconciliation (DR) is a technique used to compensate for measurement errors by means of prior knowledge of the process in the form of mathematical models so that more reliable and accurate estimates of the process states are obtained.

Data reconciliation has been widely applied to various chemical, petrochemical, and mineral processing industries during the past decades (Narasimhan et al., 2000). Typical applications of DR have been largely limited to reconciling material and energy balances, and simultaneous parameter estimation and data reconciliation for steady-state processes. However, many chemical processes are intrinsically dynamic and disturbances frequently occur. Consequently, it is desirable to develop Dynamic Data Reconciliation (DDR) strategies for dynamic processes so that at every sampling instant more accurate and reliable process data are available in real-time for control and optimization. For discrete sampled measurements, the DDR problem is typically formulated as

$$\text{Minimize} \qquad J(\hat{\mathbf{y}}_t, \hat{\mathbf{z}}_t) = \sum_{t=0}^{t_c} [(\mathbf{y}_t - \hat{\mathbf{y}}_t)^T \mathbf{V}^{-1} (\mathbf{y}_t - \hat{\mathbf{y}}_t)] \tag{4.1}$$

$$\text{subject to} \quad \mathbf{f}[\frac{d\hat{\mathbf{y}}(t)}{dt}, \hat{\mathbf{y}}(t), \hat{\mathbf{z}}(t)] = \mathbf{0}$$

$$\mathbf{g}[\hat{\mathbf{y}}(t), \hat{\mathbf{z}}(t)] = \mathbf{0}$$

$$\mathbf{h}[\hat{\mathbf{y}}(t), \hat{\mathbf{z}}(t)] \geq \mathbf{0}$$

where $\mathbf{y}_t$ is a $M \times 1$ vector of observed values of measured process variables at time t (t = 0, 1, 2, ..., t_c), $\hat{\mathbf{y}}_t$ is a $M \times 1$ vector of reconciled values for the measured process variables at time t, $\hat{\mathbf{z}}_t$ is a $N \times 1$ vector of estimates for unmeasured variables or model parameters at time t, $\mathbf{V}$ is a $M \times M$ covariance matrix of the measurements, $\mathbf{f}$ is a $C \times 1$ vector of differential model equality constraints, $\mathbf{g}$ is a $S \times 1$ vector of algebraic model equality constraints, and $\boldsymbol{h}$ is a $E \times 1$ vector of inequality model constraints including simple upper and lower bounds.

The constraints expressed by model equations $\mathbf{f}$, $\mathbf{g}$, and $\boldsymbol{h}$ in the DDR formulated by equation (4.1) can be linear or nonlinear. For linear model constraints, the optimal solutions to the problem can be recursively given by a Kalman filter (Narasimhan et al., 2000). The Kalman filter has been widely used in control of complex dynamic systems such as electronics manufacturing and the guidance of aircrafts, ships, and spacecrafts. However, its practical application in chemical engineering has been relatively infrequent. For nonlinear model constraints, an extended Kalman filter, which is based on linear approximations of nonlinear models, can be applied to solve this problem. Another approach, called Nonlinear Dynamic Data Reconciliation (NDDR), has been discussed by Liebman et al. (1992), Ramamurthi et al. (1993), Albuquerque et al. (1996) and Binder et al. (2002) to handle the nonlinear constraints in the DDR problem.

Previous work in dynamic data reconciliation has been mainly devoted to optimal estimation of process states (Liebman et al., 1992; Albuquerque et al., 1996; Barbosa et al., 2000; Soderstrom et al., 2000; Oisiovici et al.; 2000; Binder et al., 2002). Few papers have addressed the problems of using estimated state values instead of measurements for controllers, and assessing the controller performance with the DDR strategies. Ramamurthi et al. (1993) showed that dynamic data reconciliation led to better closed-loop performance, and Abu-el-zeet et al. (2002) demonstrated that NDDR, in conjunction with systematic bias detection, enhanced a model predictive control scheme. However, the degree of improvement for the controller performance was not given.

In this work, the implementation of a DDR algorithm as an integral part of conventional PID control loops (Figure 4.1) was proposed and its impact on controller performance was assessed. Two examples, a simple storage tank and a more complex distillation process, were employed in this study to illustrate the implementation of the DDR strategies. A fundamental dynamic mass balance model was used in the DDR algorithm for the storage tank, whereas empirical dynamic models obtained through model identification were used in the case of the distillation column.

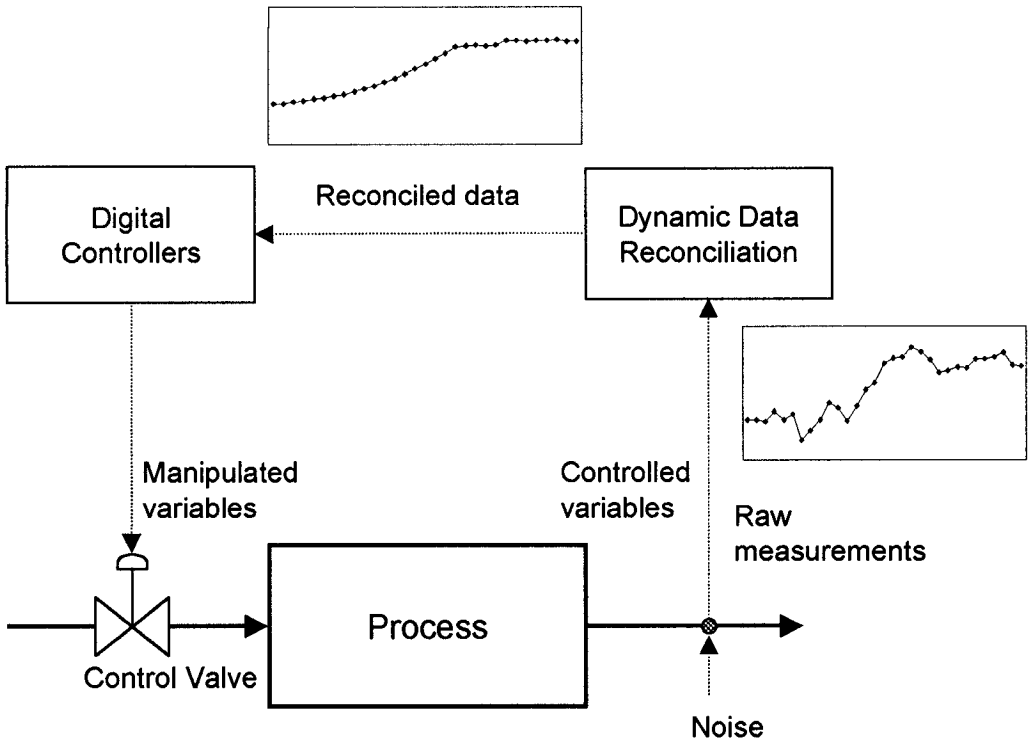

Figure 4.1 Application of dynamic data reconciliation in process control

4.2 Formulation of Dynamic Data Reconciliation Algorithm

For any process, two sources of information are usually available about the true values of the process variables. One source is a set of mathematical models representing the physical system, and the other is a set of process measurements. At each sampling time, both predicted and measured values are available. No mathematical models of a physical system are exact, nor are the measurements of variables perfect. Both contain some degree of uncertainty. It would seem reasonable that the estimates of the true values of process variables be a compromise between the measurements and the model predicted values. Under the assumption that measurement and model errors are normally distributed with mean-zero and known variances, the estimates (reconciled or adjusted

values) of the measured process variables can be obtained by solving the weighted least-squares optimization problem such that

$$\text{Minimize } J(\hat{\mathbf{y}}_t) = \left(\begin{bmatrix} \mathbf{y}_{1,t} \\ \cdots \\ \mathbf{y}_{2,t} \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{y}}_{1,t} \\ \cdots \\ \hat{\mathbf{y}}_{2,t} \end{bmatrix} \right)^T \mathbf{V}^{-1} \left(\begin{bmatrix} \mathbf{y}_{1,t} \\ \cdots \\ \mathbf{y}_{2,t} \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{y}}_{1,t} \\ \cdots \\ \hat{\mathbf{y}}_{2,t} \end{bmatrix} \right) + (\mathbf{y}_{\mathrm{m},1,t} - \hat{\mathbf{y}}_{1,t})^T \boldsymbol{\Omega}^{-1} (\mathbf{y}_{\mathrm{m},1,t} - \hat{\mathbf{y}}_{1,t})$$

$$\text{subject to } \mathbf{y}_l \le \hat{\mathbf{y}}_t \le \mathbf{y}_u \tag{4.2}$$

where the $M \times 1$ vector of measurements, $\mathbf{y}_t$, for the M process variables at time t, is partitioned into $\mathbf{y}_{1,t}$ and $\mathbf{y}_{2,t}$. The $\mathbf{y}_{1,t}$ is a $M_1 \times 1$ vector of measurements for M_1 process variables whose values can be predicted by process models, and the $\mathbf{y}_{2,t}$ is a $M_2 \times 1$ vector of measurements for M_2 independent variables. $\hat{\mathbf{y}}_{1,t}$ and $\hat{\mathbf{y}}_{2,t}$ are the vectors of reconciled values for the measurements, $\mathbf{y}_{1,t}$ and $\mathbf{y}_{2,t}$, respectively. $\mathbf{y}_{m,1,t}$ is a $M_1 \times 1$ vector of model predicted values at time t. $\mathbf{y}_l$ and $\mathbf{y}_u$ are $M \times 1$ vectors of lower and upper bounds for process variables. $\mathbf{V}$ is a $M \times M$ variance-covariance matrix of the measurements, and $\boldsymbol{\Omega}$ is a $M_1 \times M_1$ variance-covariance matrix of the model predictions.

The model predicted values, $y_{m,1,t}$ in $\mathbf{y}_{m,1,t}$, can be linear or nonlinear functions of variables in $\hat{\mathbf{y}}_t$, or can be functions of other known process variables. The solution to the optimization problem of equation (4.2) can be obtained by a minimization routine such as the quasi-Newton algorithm.

In equation (4.2), the variances and covariances of measurements in $\mathbf{V}$ can be obtained from raw measurements. However, it is sometimes difficult to obtain the variances and covariances of the model predicted values in $\mathbf{\Omega}$. This is not only because of model inadequacies, but also because of various kinds of unknown disturbances corrupting the process. As a result, we have treated the ratio of the variance of the model predictions, σ_m^2, to that of the process measurements, σ^2, as tuning parameters in the DDR algorithm.

When σ_m^2/σ^2 is used as a tuning parameter for the DDR algorithm, the locus of the reconciled value, $\hat{y}_t$, typically follows an asymptotic curve as a function of σ_m^2/σ^2. Figure 4.2, for example, illustrates the locus of a reconciled value with change in σ_m^2/σ^2 for one measurement of a mono-variable. The measured value for the liquid level of a tank is 60.0 cm, and the model predicted value of the liquid level is 60.5 cm, at time *t*. The reconciled value is a point on the curve whose location depends on the values of σ_m^2/σ^2. The reconciled value is bounded by the measured and predicted values. If the variance of the model predicted value is given a relatively small value compared to the variance of the measurements, the reconciled data tend to be close to the model predicted values. On the other hand, for a relatively large value, the reconciled data tend to be close to the measurements. There are two limiting cases: i) $\sigma_m^2/\sigma^2 = 0$ corresponds to an exact model, and the reconciled value is equal to the model predicted value; ii) $\sigma_m^2/\sigma^2 \to \infty$ indicates that the level of confidence on model predictions is poor, and the reconciled value approaches the measured value.

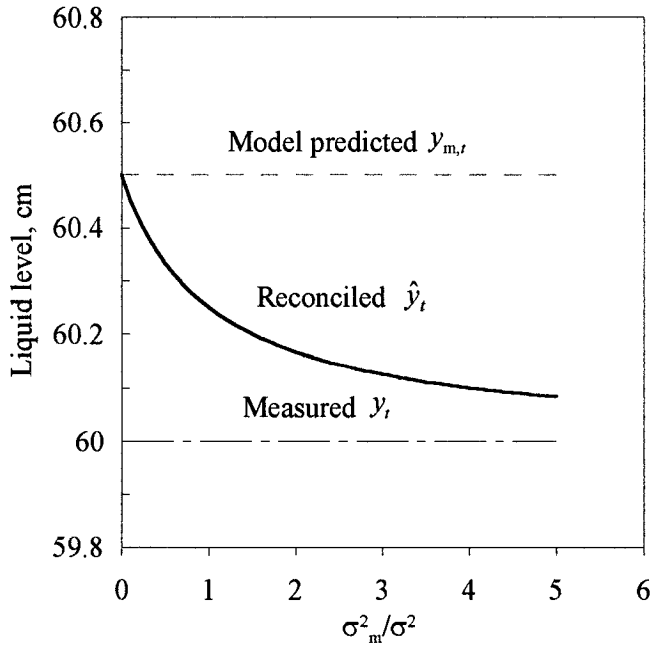

Figure 4.2 Locus of reconciled data as a function of the ratio of model to measurement uncertainty in the DDR algorithm

It should be noted that reliable process models play a key role for good performance of the DDR algorithm. The process models can be fundamental models that are based on mass and heat balances, thermodynamics, kinetics, etc. However, for complex processes, it is normally not practical to develop complex mechanistic process models. Consequently, simple input-output empirical models can often be employed to capture the essence of the dynamics of the process.

4.3 Simulation Examples

The dynamic data reconciliation strategy developed in Section 4.2 was tested using simulations of (i) simple level control of a liquid storage tank, (ii) multivariable control of a binary distillation column.

4.3.1 Storage Tank

The schematic diagram of the storage tank process is shown in Figure 4.3. The diameter and the height of the tank were 1.0 m and 1.2 m, respectively. A PI controller was used to regulate the liquid level of the tank by manipulating the outlet flow. The feed flow to the tank was measured but not controlled. The sampling interval was 1 minute. Under steady state, the two measured variables, the feed flow rate and the liquid level were 1.8 m^3/h and 0.6 m, respectively. The measurement errors were assumed to be Gaussian white noise, and the standard deviations for the measurements of the feed flow rate and the liquid level were set at 0.036 m^3/h and 0.012 m, respectively. The dynamics of the measuring device and control valve were neglected.

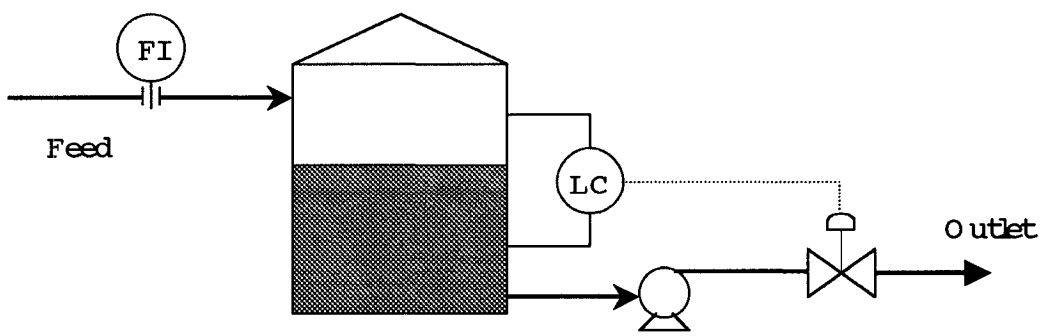

Figure 4.3 Schematic diagram of a storage tank process

Values of the controller parameters were obtained by subjecting the process to a 20% step change in the feed flow rate and then minimizing the following objective function

$$\Phi\ (\mathbf{K}_C, \tau_I) = \sum_{i=1}^{k} \left(\alpha_i \times ISE_i + \beta_i \times ISDU_i \right) \tag{4.3}$$

where Φ is the cost function to evaluate controller performance. $\mathbf{K}_C$ and τ_I are $k \times 1$ vectors of controller parameters, and k is the total number of control loops. ISE_i is the integral of the squared error of the controlled variable of controller i, $ISDU_i$ is the integral of squared differences of the values of manipulated variable between sampling times t and $t-1$ of controller i, α_i and β_i are the weighting factors. The minimization of equation (4.3) was carried out by a quasi-Newton method. The initial optimal values of the controller parameters were $K_C = -1.87$ $m^2.h^{-1}$ and $\tau_I = 30.0$ min. It is important to note that these values of the control parameters were obtained by considering noisy measurements. The performance of the controller was tested with the same feed perturbation and results of this simulation are presented in Figure 4.4. The controlled variable, liquid level, returns to its setpoint in about 180 minutes. The manipulated variable, the outlet flow, exhibits considerable variation because the noisy measurements affected the calculated control moves. The values of ISE, $ISDU$ and the overall performance measure for the controller, Φ, are presented as Case I of Table 4.1.

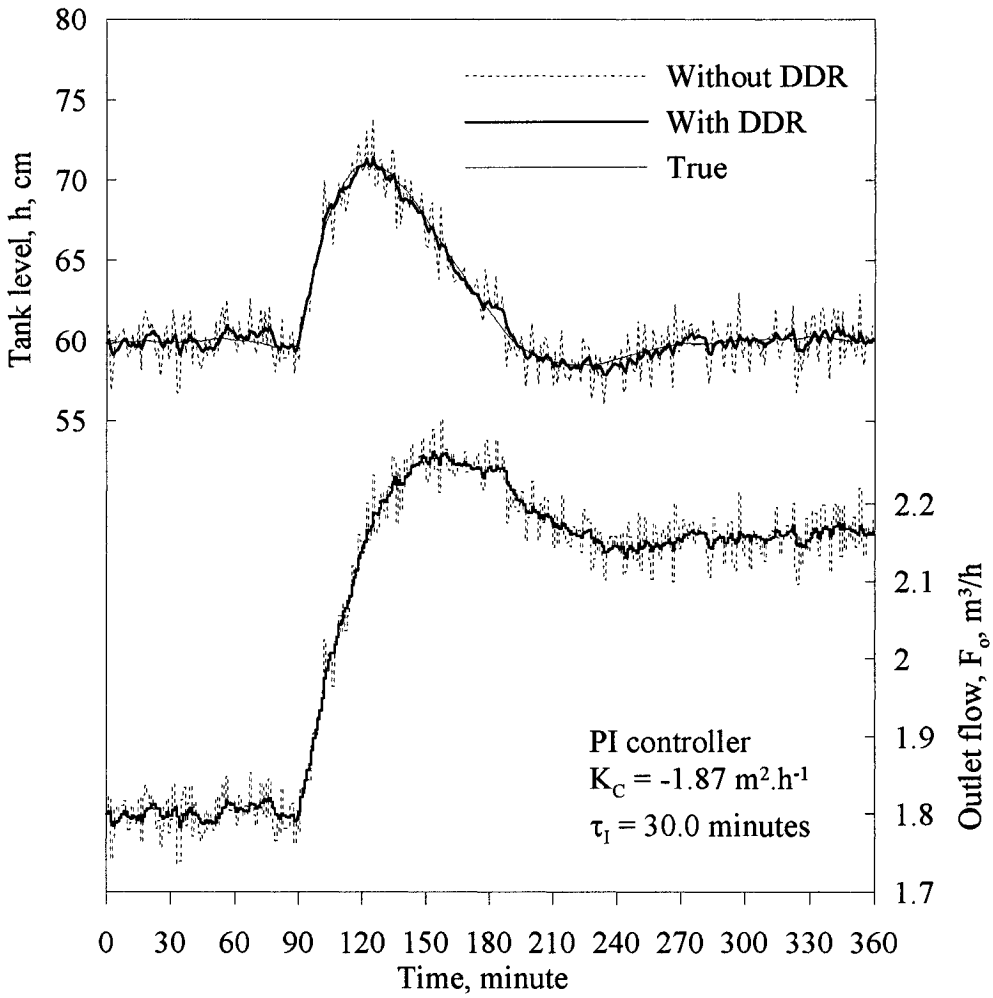

Figure 4.4 Performance of controller with initial optimized parameters, for a 20% step increase in feed flow rate at t=90 min

Table 4.1 Comparison of level controller performance for the storage tank

Case	DDR	ISE (α=0.3)	ISDU (β=0.7)	$\Phi(K_C, \tau_I)$
I (K_C = -1.87, τ_I =30)	No	3.38	1.37	1.97
II (K_C = -1.87, τ_I =30)	Yes	3.38	0.03	1.03
III (K_C = - 4.63, τ_I =30)	Yes	0.59	0.23	0.34

To attenuate the impact of measurement noise, a DDR algorithm was embedded into the control loop to reconcile the liquid level before calculating the control moves. For this simple process, equation (4.2) can be explicitly written in the following form:

$$\text{Minimize } J(\hat{h}_t, \hat{F}_{i,t-1}) = \frac{1}{\sigma_h^2}(h_t - \hat{h}_t)^2 + \frac{1}{\sigma_{F_i}^2}(F_{i,t-1} - \hat{F}_{i,t-1})^2 + \frac{1}{\sigma_m^2}(h_{m,t} - \hat{h}_t)^2 \qquad (4.4)$$

$$\text{subject to } 0 < \hat{h}_t < h_{\max}$$

$$0 < \hat{F}_{i,t-1} < F_{\max}$$

where h_t, and $F_{i,t-1}$ are the measured values for the liquid level and feed flow, respectively. $\hat{h}_t$ and $\hat{F}_{i,t-1}$ are the reconciled values for the liquid level and feed flow, respectively. $h_{\max}$ and $F_{\max}$ are the upper bounds of the two variables. $h_{m,t}$ is the model predicted value for the liquid level at time t, which is given by the mass balance in the discrete form as

$$h_{m,t} = \hat{h}_{t-1} + \frac{\Delta t}{A}(\hat{F}_{i,t-1} - F_{o,t-1}) \qquad (4.5)$$

where $F_{o,t-1}$ is the value of outlet flow given by the controller at time t, Δt is the sampling time interval, and A is the cross-sectional area of the tank.

The ratio of the variance of the model predicted values to the variance of the measurements was set at $\sigma_m^2/\sigma^2 = 0.25$ in testing the performance of the DDR algorithm. The same feed flow perturbation was used to evaluate the controller performance with the DDR algorithm embedded in the control loop. The simulation results are also presented in Figure 4.4. In this case, the reconciled data for the liquid level were less noisy than the raw measurements and close to their true values. The manipulated variable displayed much smaller variations when the reconciled data were used to calculate the control action. The corresponding values of *ISE*, *ISDU* and Φ are listed as Case II of Table 4.1. The value of *ISE* remained constant, whereas the value of *ISDU* decreased significantly compared with Case I. The cost function of the overall performance, Φ, dropped to 1.03 from 1.97, which is a 48% reduction.

With the same cost function of the controller performance and the same external perturbation of the feed flow, the controller parameters were optimized again with the DDR algorithm embedded inside the control loop. The new optimized parameters were $K_C = -4.63$ $m^2.h^{-1}$ and $\tau_I = 30.0$ min indicating an increase in the controller gain by a factor of 2.5 and no change in the integral time. The results of this simulation are presented in Figure 4.5. The response of the process became faster, and the liquid level returned to its setpoint in about 100 minutes compared to 180 minutes for Cases I and II. In addition the maximum deviation from setpoint decreased from approximately 10.0 cm to 5.0 cm. The variations of the manipulated variable increased slightly but were judged acceptable considering the more aggressive controller. The values of *ISE*, *ISDU* and Φ are summarized as Case III of Table 4.1. The *ISE* was reduced. The *ISDU* increased

compared with Case II, but was smaller than that of Case I. The cost function of the controller reduced to 0.34, a 83% reduction compared with 1.97 for Case I and a 67% reduction compared with 1.03 for Case II.

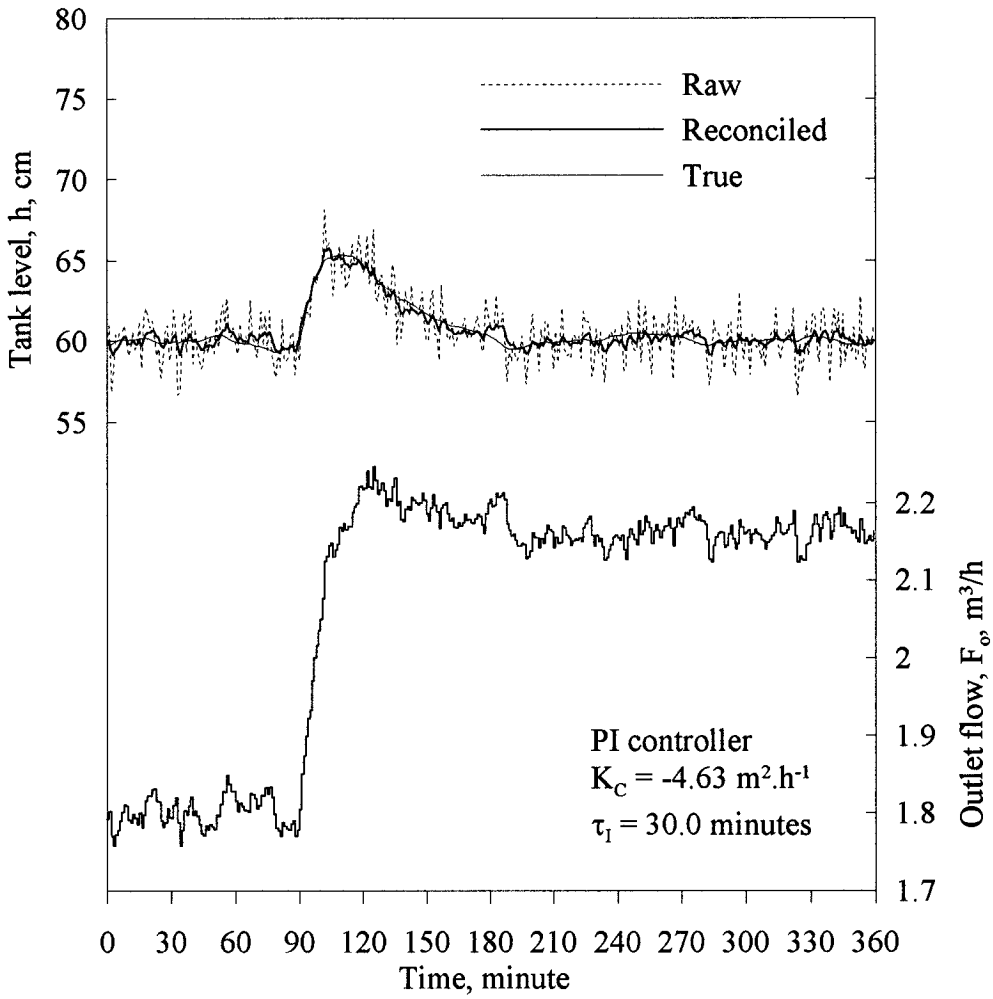

Figure 4.5 Performance of controller with final optimized parameters, for a 20% step increase of feed flow rate at t=90 min

This simple example demonstrates that much better control performance can be achieved by embedding a DDR algorithm inside a feedback control loop. This better performance was obtained because DDR combined information from both process measurements and a process model to provide a better estimate (more precise and consistent) of the level at each sampling time which decreased excessive manipulation and allowed more aggressive control action to be taken. The better estimate was a compromise between the measured and the predicted values of the controlled variable as shown in Figure 4.2.

4.3.2 Distillation Column

A similar DDR algorithm was implemented for a more complex process, a binary (benzene/toluene) distillation column, and the performance of the DDR was evaluated. The schematic diagram of the distillation is shown in Figure 4.6. Two inferential feedback PI-controllers, TIC-D and TIC-B, were used to control the top and bottom compositions of benzene by manipulating the reflux flow rate and the reboiler heat duty, respectively. Two other feedback PI-controllers, LIC-D and LIC-B, were used to control the reflux drum and column base liquid levels by manipulating the distillate flow rate and the bottom product flow rate, respectively. The discrete sampling time was 30 seconds for all control loops. The external disturbances were the feed flow rate and feed composition. The nominal steady-state values for all measurements and their typical noise levels are listed in Table 4.2.

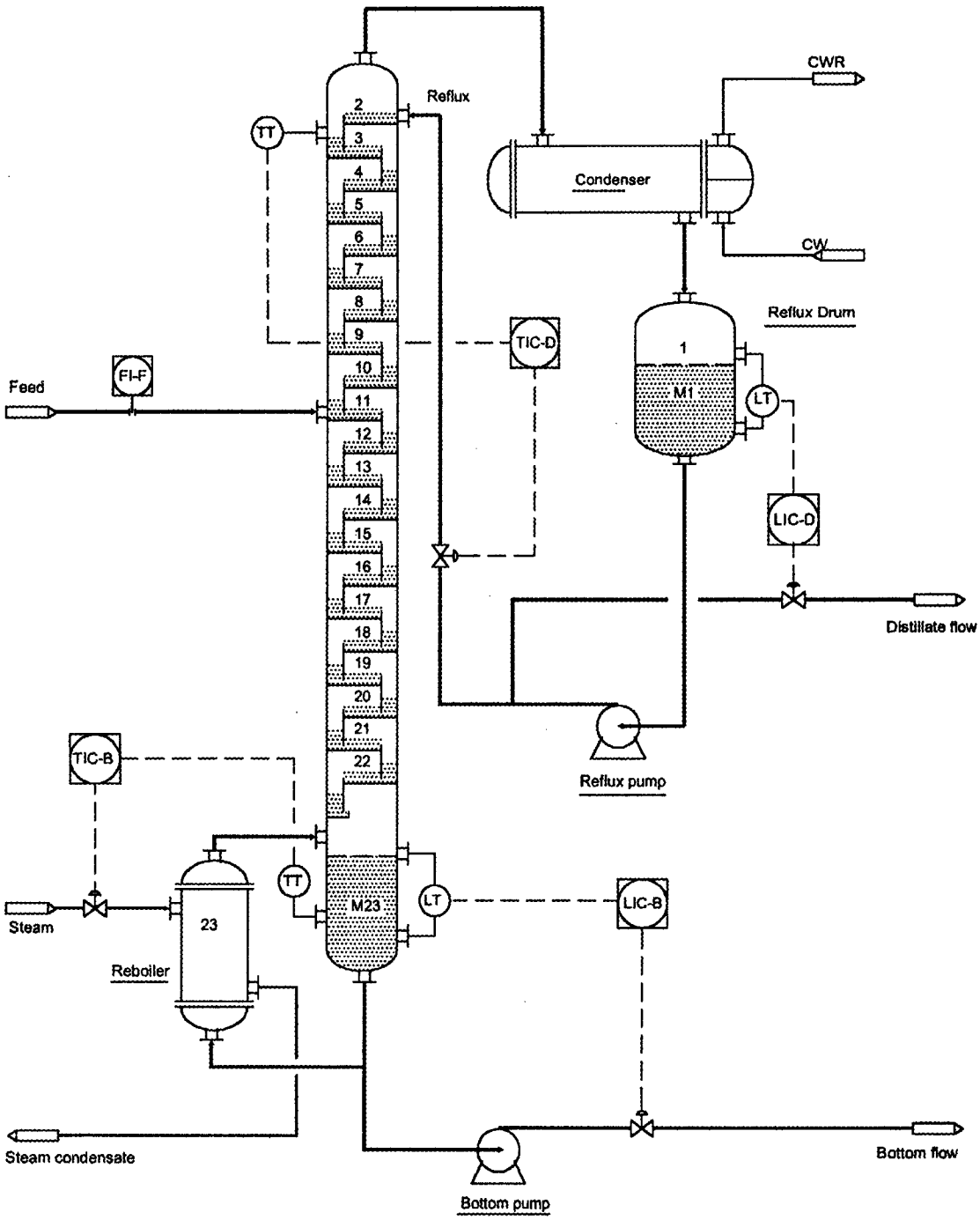

Figure 4.6 Schematic diagram of the binary distillation column

Table 4.2 Nominal steady-state values and noise levels of measured variables

Variable	Units	Steady-state values	Standard deviation
Feed flow (F_F)	kmol/h	100.0	0.50
Column top temperature (T_D)	°C	84.2	0.25
Column bottom temperature (T_B)	°C	117.4	0.25
Reflux drum level (H_D)	m	0.50	0.02
Column base level (H_B)	m	0.70	0.02

A simulator of this distillation column was developed for this study. The simulator was based on rigorous distillation dynamic models, except for the following assumptions: 1) liquid holdups on trays, downcomers, reflux drum, and column base were considered to be well-mixed continuous stirred tanks; 2) vapor hold up was neglected; 3) tray efficiency was assumed to be 100%; 4) dynamics of the condenser, reboiler, measuring devices, and control valves were negligible compared with the dynamics of the distillation column; 5) measurement noise was assumed to be Gaussian white noise.

The tuning parameters of the four PI controllers were optimized by minimizing the cost function defined by equation (4.3) with a 20% step change in the feed flow rate. The controller parameters that were obtained are listed as Case A of Table 4.3, and the results of the controller performance are presented in Figure 4.7. The column base liquid level and the bottom temperature were affected more rapidly, whereas the responses of reflux drum level and the top temperature were delayed and changed with relatively slow dynamics. The values of *ISE* and *ISDU* for each control loop and the overall cost function of the control system are summarized as Case I of Table 4.4.

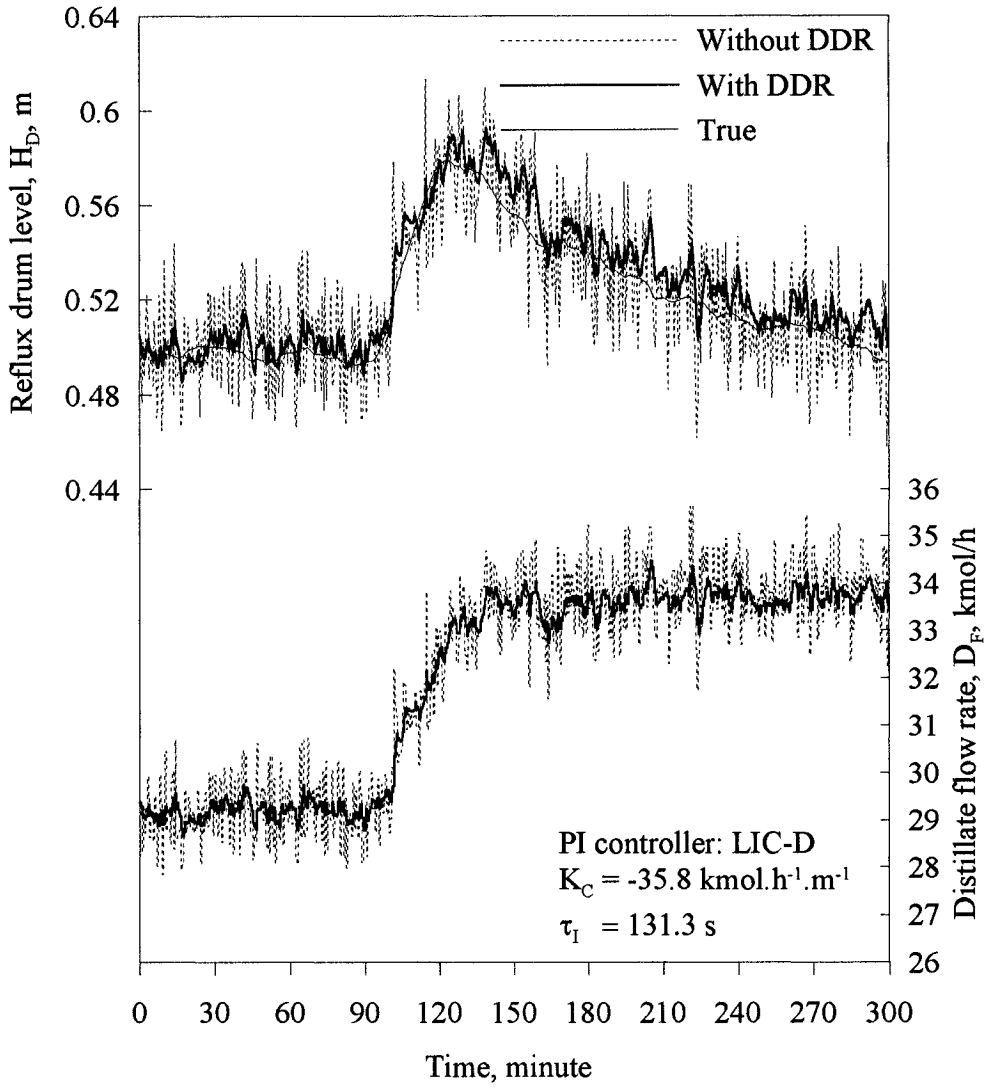

Figure 4.7(a) Performance of reflux drum level controller with initial optimized parameters, for a 20% step increase in feed flow rate

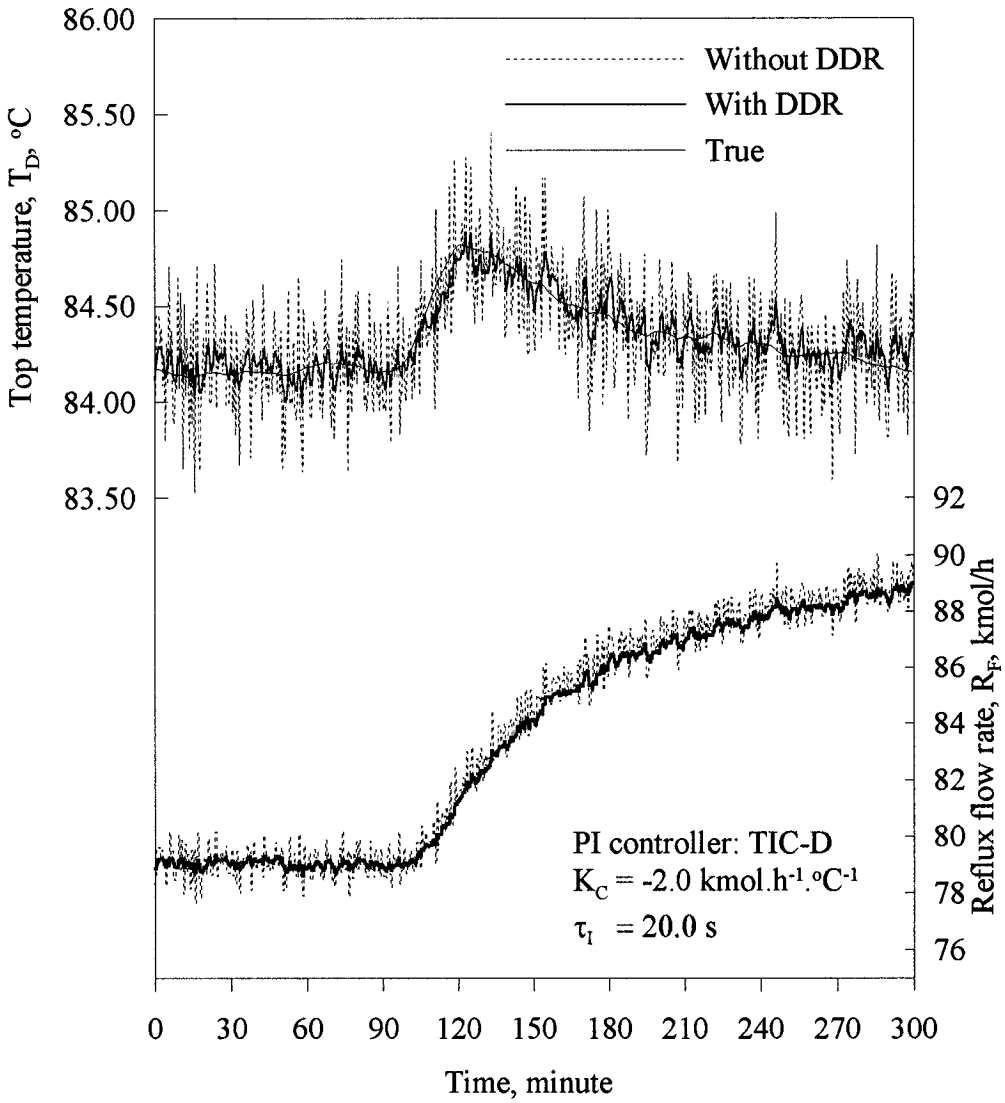

Figure 4.7(b) Performance of top temperature controller with initial optimized parameters, for a 20% step increase on feed flow rate

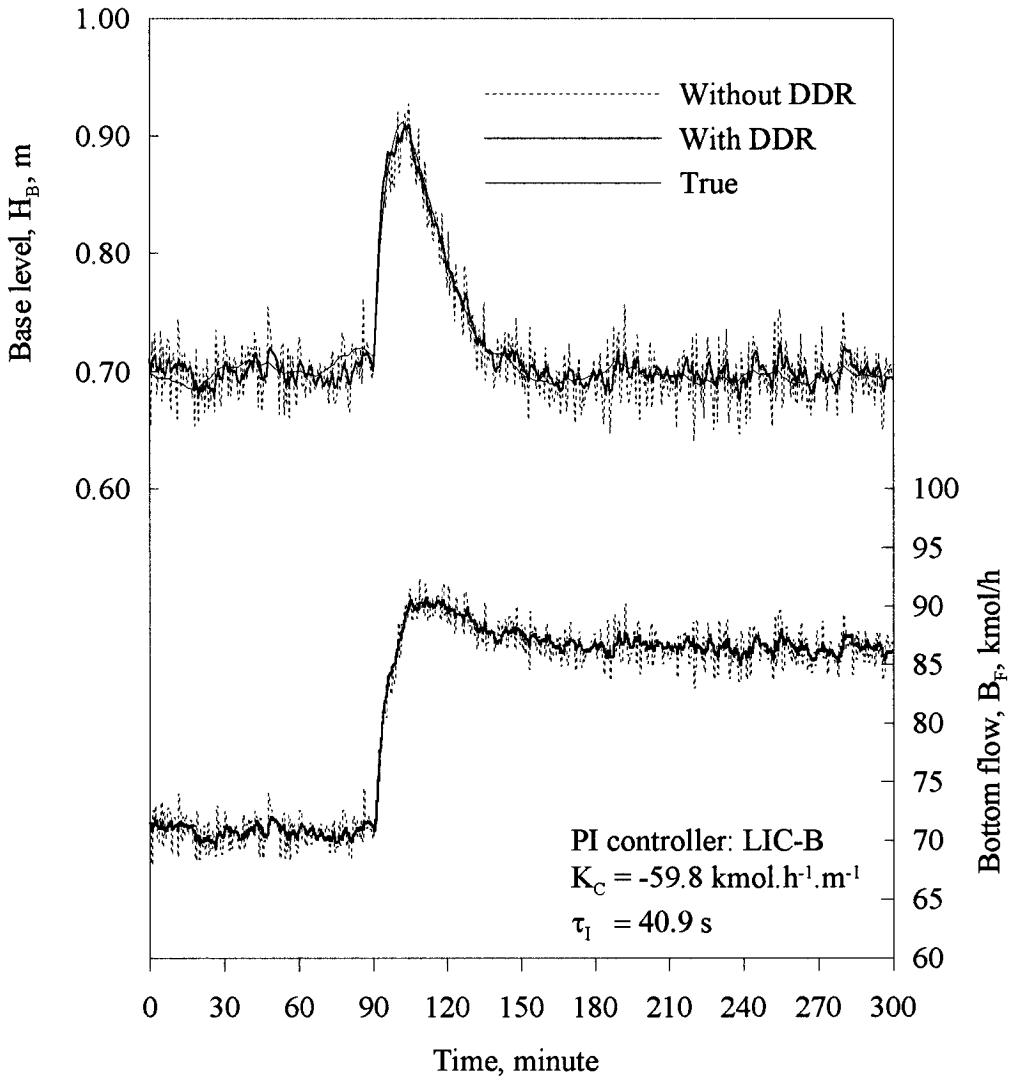

Figure 4.7(c) Performance of column base level controller with initial optimized parameters, for a 20% step increase in feed flow rate

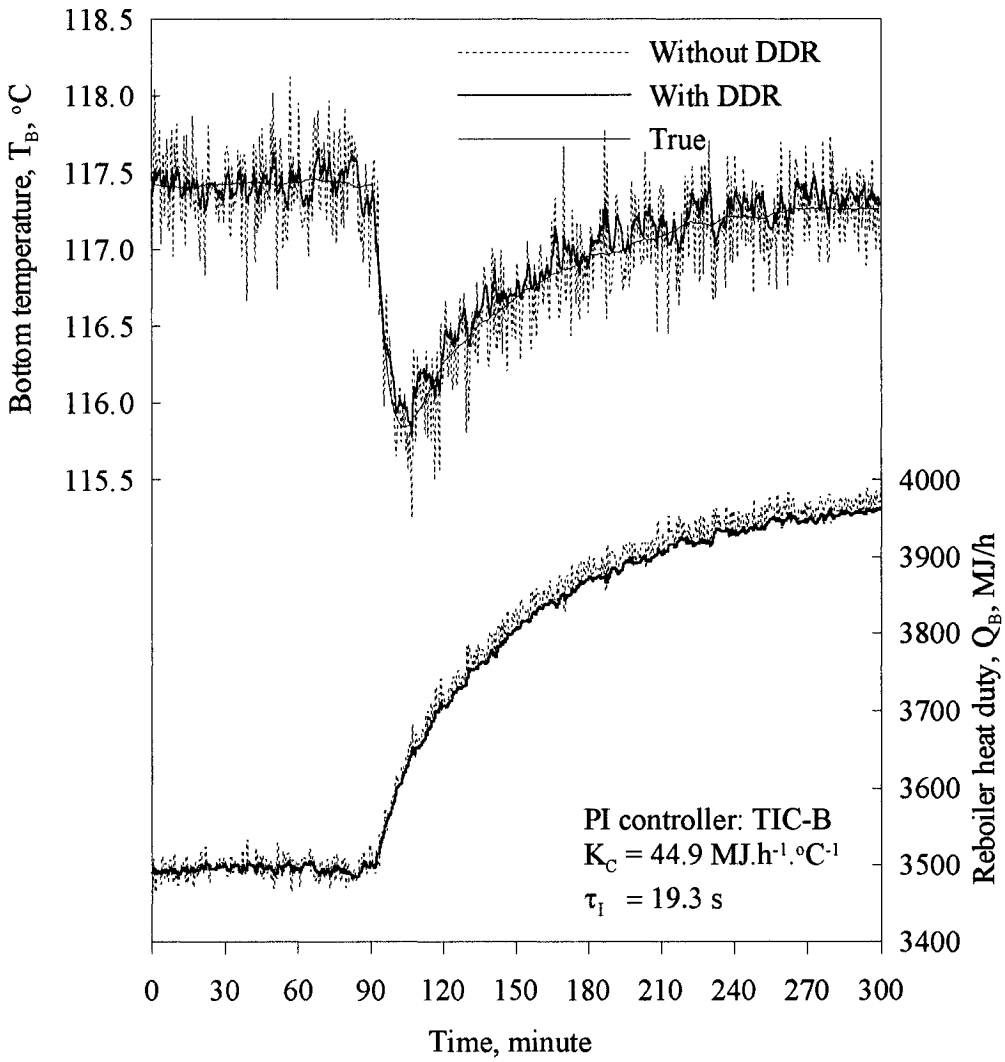

Figure 4.7(d) Performance of bottom temperature controller with initial optimized parameters, for a 20% step increase in feed flow rate

Table 4.3 Controller tuning parameters for the distillation column

Set of tuning parameters	LIC-D		TIC-D		LIC-B		TIC-B	
	K_C $kmol.h^{-1}.m^{-1}$	τ_I s	K_C $kmol.h^{-1}.°C^{-1}$	τ_I s	K_C $kmol.h^{-1}.m^{-1}$	τ_I s	K_C $MJ.h^{-1}.°C^{-1}$	τ_I s
A	-35.8	131.3	-2.0	20.0	-59.8	40.9	44.9	19.3
B	-74.9	14.1	-5.8	10.1	-192.5	14.7	142.1	10.7

Table 4.4 Comparison of controller performance for the distillation column

Case	Tuning set	DDR	LIC-D α=0.1, β=0.1		TIC-D α=0.2, β=0.1		LIC-B α=0.05,β=0.05		TIC-B α=0.3, β=0.1		Φ
			ISE	ISDU	ISE	ISDU	ISE	ISDU	ISE	ISDU	
I	A	No	0.89	0.63	0.65	0.37	0.39	1.00	0.97	0.48	0.73
II	A	Yes	0.64	0.02	0.70	0.01	0.39	0.04	1.10	0.02	0.56
III	B	Yes	0.08	0.10	0.18	0.11	0.02	0.51	0.16	0.18	0.16

Figure 4.7 shows that the responses of the column to the feed flow disturbance are slow, and the manipulated variables have relatively large variations because of the noisy measurements. To reduce the propagation of the measurement noise through the control loops, a DDR algorithm was developed for the distillation column. Because of the complexity of the distillation process, it is not practical to develop or use fundamental models for predictions of the controlled variables in the implementation of the DDR algorithm. Instead, empirical models were identified and used in this work.

The open-loop process reaction curve for a pair of input/output variables with a step change in the input was simulated. Then, pure-integration-plus-dead-time and first-order-plus-dead-time models were used to approximate the process responses. The model parameters, K_p, θ, and τ_p, were obtained by fitting the simulated data using least-squares technique (Marlin, 2000). The identified empirical models for the distillation column, in the Laplace domain, are given by

$$\begin{bmatrix} H_D(s) \\ T_D(s) \\ H_B(s) \\ T_B(s) \end{bmatrix} = \mathbf{G}_U(s) \begin{bmatrix} D_F(s) \\ R_F(s) \\ B_F(s) \\ Q_B(s) \end{bmatrix} + \mathbf{G}_D(s) \begin{bmatrix} F_F(s) \\ X_F(s) \end{bmatrix} \tag{4.6}$$

where H_D is the reflux drum liquid level, T_D is the top temperature, H_B is the column base liquid level, T_B is the reflux base temperature, D_F is the distillate flow, R_F is the reflux flow, B_F is the bottom flow, Q_B is the reboiler heat duty, F_F is the feed flow, X_F is the feed composition. In equation (4.6),

$$\mathrm{G}_U(s) = \begin{bmatrix} \dfrac{-0.1087}{s} & \dfrac{-0.1087}{s} & 0 & \dfrac{0.0038e^{-10s}}{s} \\ 0 & \dfrac{-0.0621e^{-14s}}{176s+1} & 0 & \dfrac{0.0297e^{-120s}}{s} \\ 0 & \dfrac{0.1228e^{-30s}}{s} & \dfrac{-0.1382}{s} & \dfrac{-0.0037e^{-10s}}{s} \\ 0 & \dfrac{-1.8344e^{-218s}}{3672s+1} & 0 & \dfrac{0.0031e^{-28s}}{126s+1} \end{bmatrix}$$

is the matrix of transfer functions related to the manipulated variables, and

$$G_D(s) = \begin{bmatrix} 0 & \dfrac{-9.0e^{-20s}}{s} \\ \dfrac{-0.0426e^{-20s}}{s} & \dfrac{-18.88e^{-95s}}{372s+1} \\ \dfrac{0.1429e^{-10s}}{s} & \dfrac{9.1e^{-40s}}{s} \\ \dfrac{-0.1335e^{-60s}}{513s+1} & \dfrac{-114.7e^{-150s}}{917s+1} \end{bmatrix}$$

is the matrix of transfer functions related to external disturbances.

The empirical models in equation (4.6) were discretized by z-transformation. Then, the discrete dynamic models were used in the prediction of the values of the four controlled variables at each sampling time t in the DDR algorithm. The performance of the DDR algorithm developed for the column was tested by setting $\sigma_m^2/\sigma^2 = 0.3$ for each controlled variable, and the model predicted values of the controlled variables were assumed uncorrelated. The performance of the controllers along with the embedded DDR algorithm were again evaluated with the same feed flow perturbation, and the simulation results are also presented in Figure 4.7.

The reconciled data are significantly less noisy than the raw measurements, and the manipulated variables display relatively smaller variations with the DDR. The values of

ISE and *ISDU* for each loop and the overall cost function of the controllers, Φ, are summarized as Case II of Table 4.4. The *ISE* of the control loop LIC-D was reduced by 28%, while the *ISE* of the control loop LIC-B remained constant. But the *ISE* for the control loops TIC-D and TIC-B was increased by 8% and 12%, respectively. However, the values of *ISDU* for all control loops were reduced significantly, approximately by one order of magnitude. The overall cost function of the control system, Φ, dropped to 0.56, which corresponds to a 23% reduction compared with Case I. Although there were significant reductions in *ISDU*, the response of the process remained sluggish, the time to return to setpoints remained essentially unchanged. However, the large reduction in ISDU offers an opportunity to increase the speed of response by using more aggressive control actions.

Consequently, embedded with the DDR algorithm in each control loop, the optimal controller parameters were optimized again by minimizing the same objective function with the same feed flow perturbation. The new optimized controller parameters are presented as Case B of Table 4.3. The controller gains of all controllers were increased by a factor of 2.1~3.2, while the integral times were decreased by a factor of 1.8~13.0, leading to more aggressive control. The performance of the newly optimized controllers along with the DDR algorithm was evaluated with the same perturbation. The simulation results are presented in Figure 4.8. The responses of all controlled variables were faster, while the variations of the manipulated variables were acceptable even though the controller gains were increased. The quantitative values of *ISE* and *ISDU* for each control loop and the overall cost function of the controllers, Φ, were calculated and are

listed as Case III of Table 4.4. The overall cost function of the controllers, Φ, was reduced to 0.16, which is a 78% reduction compared to Case I, and a 70% reduction compared to Case II.

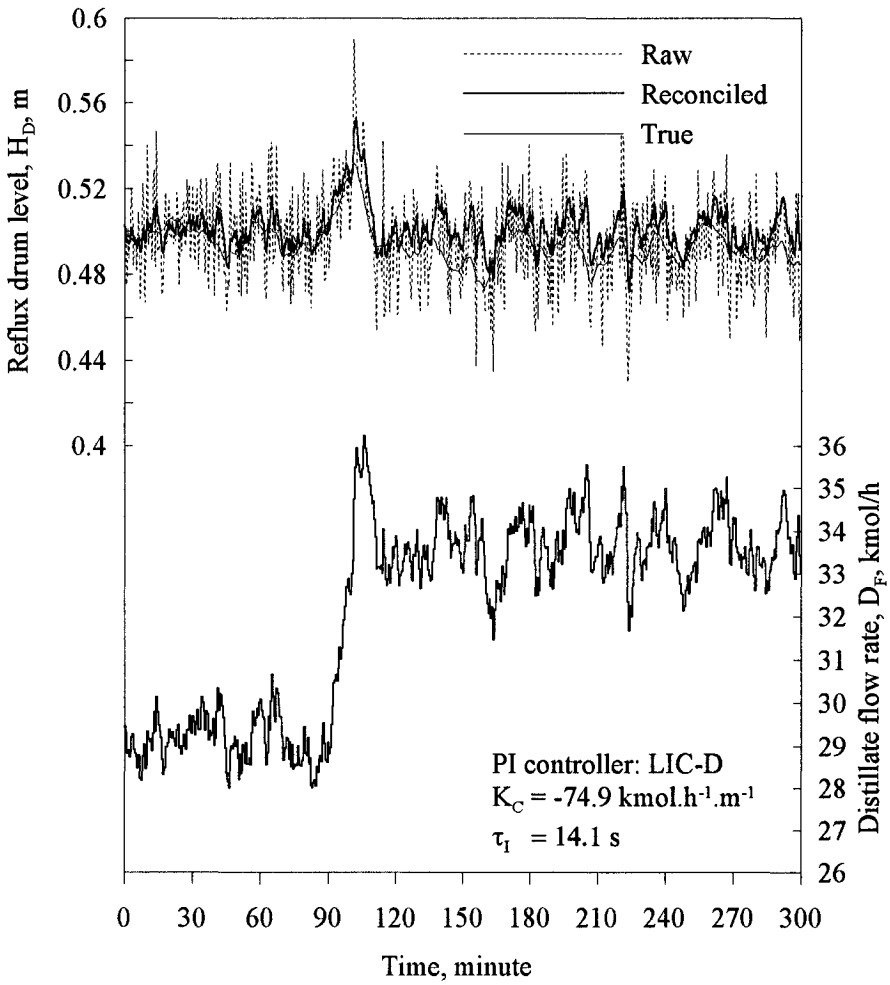

Figure 4.8(a) Performance of reflux drum level controller with final optimized parameters with DDR, for a 20% step increase in feed flow rate

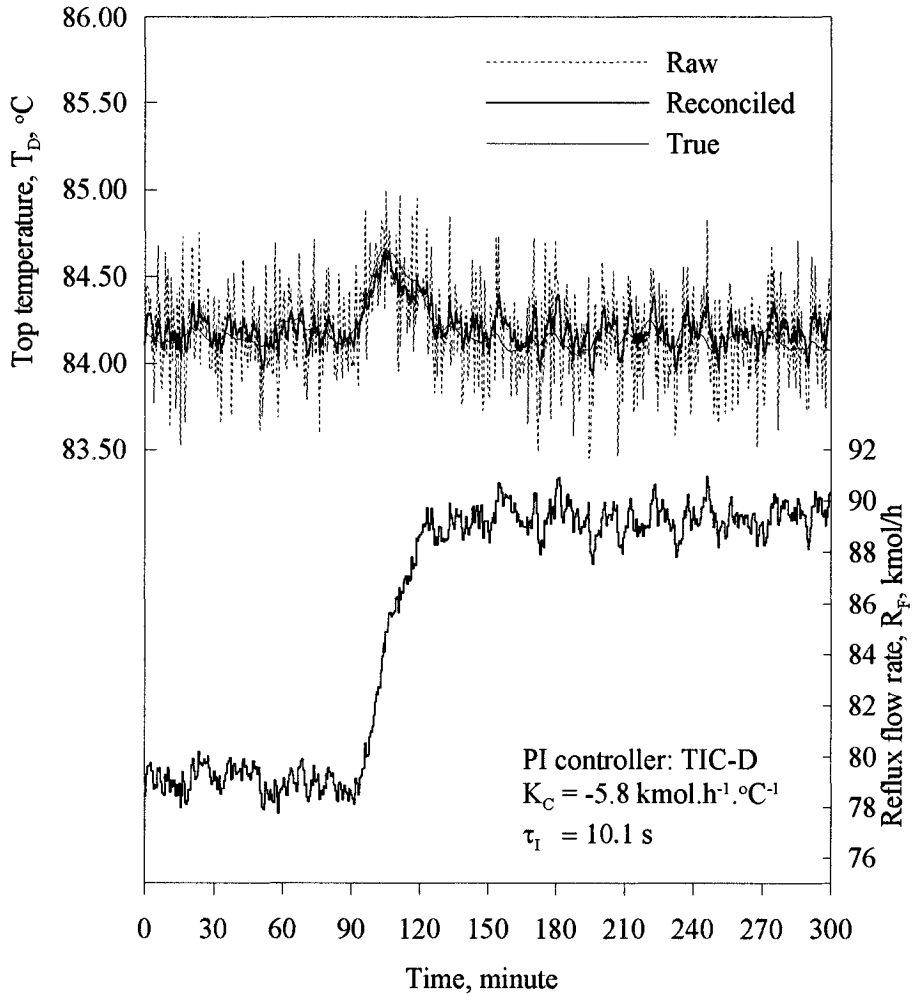

Figure 4.8(b) Performance of top temperature controller with final optimized parameters with DDR, for a 20% step increase in feed flow rate

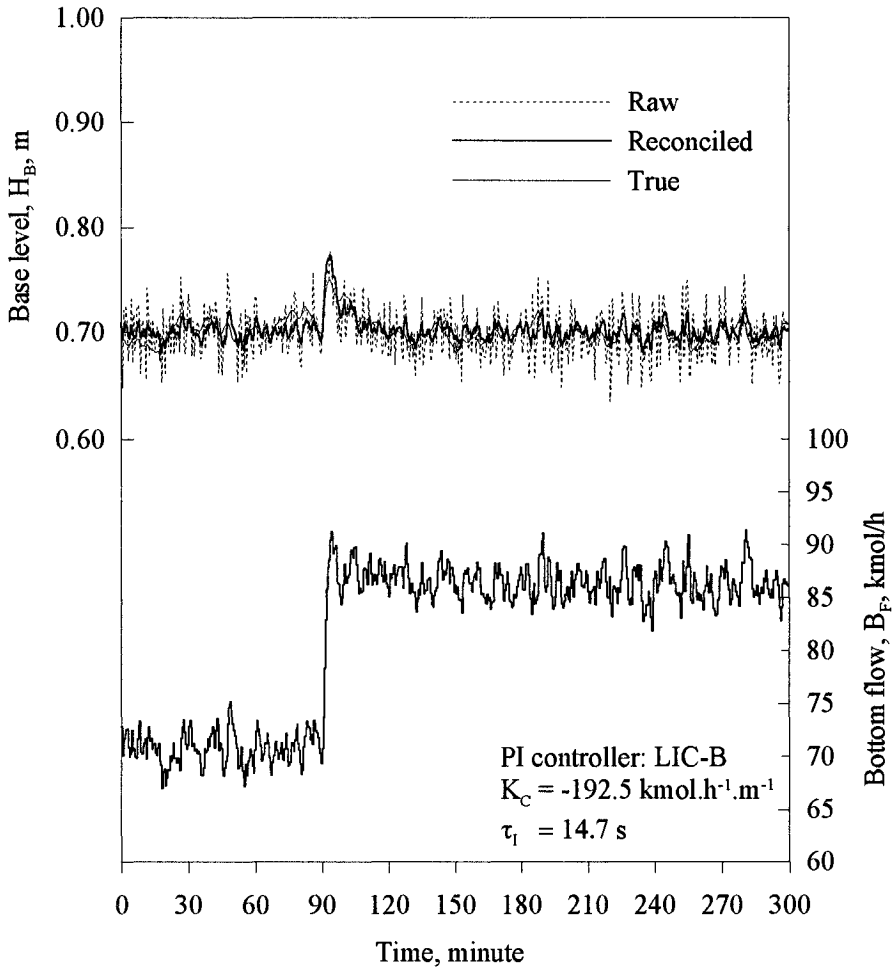

Figure 4.8(c) Performance of column base level controller with finial optimized parameters with DDR, for a 20% step increase in feed flow rate

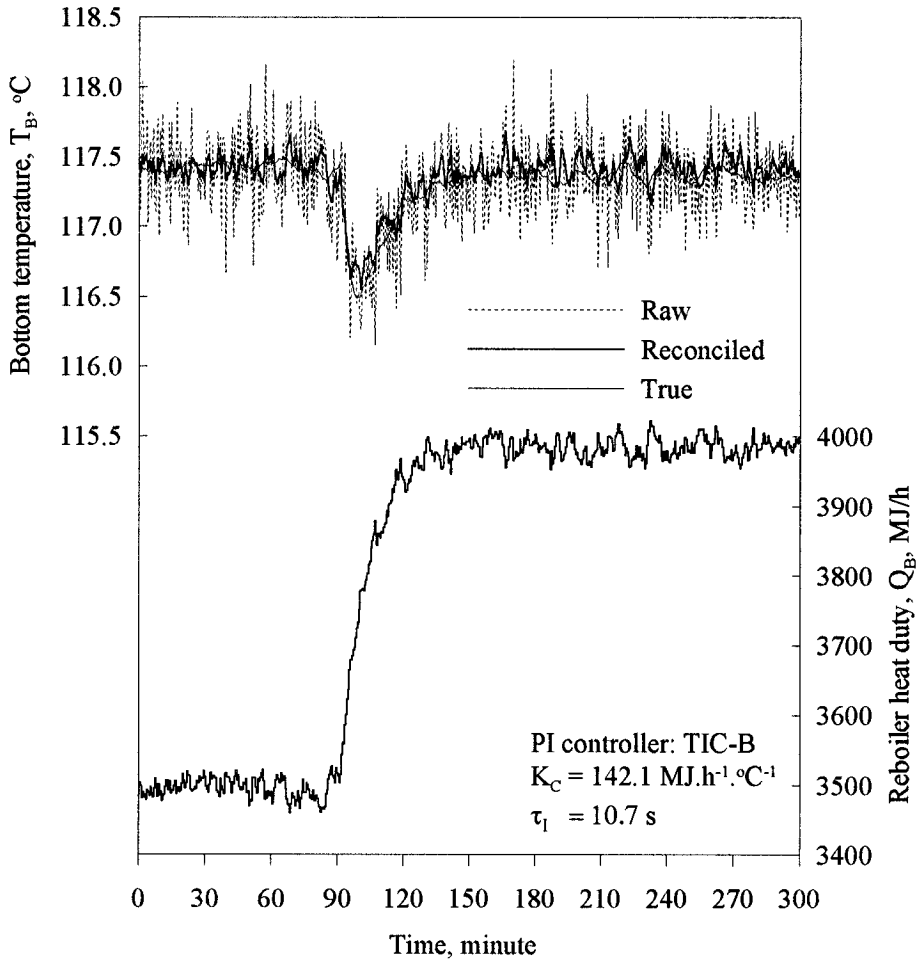

Figure 4.8(d) Performance of bottom temperature controller with final optimized parameters with DDR, for a 20% step increase in feed flow rate

Results obtained with the more complex example clearly show the benefits of using a dynamic data reconciliation to enhance the performance of controllers. Even though very simple empirical models were used to calculate the predicted values necessary for the DDR algorithm, appreciable improvement was observed. Further improvement could be achieved if better models would be used or if the reconciled data would be used to decouple the interactions between different control loops.

4.4 Conclusion

The DDR algorithm developed in this work showed be an efficient tool that can be implemented in real-time to reduce measurement noise before calculating the control action. The application of DDR results in significantly better controller performance because of the significant attenuation of measurement noise and because more aggressive controllers are allowed. It makes the controllers more aggressive, and at the same time, reduces the high-frequency variations of manipulated variables. The two processes considered in this investigation revealed that the effectiveness and efficiency of the DDR algorithm depend on the level of the model accuracy. More accurate models, that is more detailed knowledge of the process, would result in better performance of the DDR algorithm.

Nomenclature

A: Cross-sectional area, m^2

B_F: Distillation bottom flow rate, kmol/h

D_F: Distillate flow rate, kmol/h

F_F: Feed flow for distillation column, kmol/h

F_i: Feed flow for storage tank, m^3/h

F_o: Outlet flow for storage tank, m^3/h

h: Liquid level of storage tank, cm

H_D: Reflux drum liquid level of distillation column, m

H_B: Column base liquid level, m

$\mathbf{K_C}$: $k \times 1$ vector of the controller gains

Q_B: Distillation reboiler heat duty, MJ/h

R_F: Reflux flow rate of distillation, kmol/h

Δt: Discrete sampling time, s

T_D: Top temperature of distillation column, °C

T_B: Bottom temperature of distillation column, °C

$\mathbf{V}$: $M \times M$ covariance matrix of the measurements

X_F: Feed composition, mole fraction

$\mathbf{y}$: $M \times 1$ vector of raw measurements

$\mathbf{y}_{1,m}$: $M_1 \times 1$ vector of model predicted values

$\mathbf{z}$: $N \times 1$ vector of estimates for unmeasured variables

Greek letters

α_i, β_i: Weighting factors

σ^2: Variance of raw measurements

σ_h^2: Variance of measurement for tank level

$\sigma_{F_i}^2$: Variance of measurement for tank feed flow

σ_m^2: Variance of model predicted values

Φ: Cost function of controller performance

Ω: $M_1 \times M_1$ covariance matrix of model predicted values

τ_I: $k \times 1$ vector of the controller integral time

Superscripts and Subscripts

l: Lower bound

u: Upper bound

t: Time instant

$\wedge$: Estimate

Acronyms

ISE: Integral of squared errors for controlled variable

ISDU: Integral of squared difference of manipulated variable

References

Abu-el-zeet, Z.H.; P.D. Roberts; and V.M. Becerra, "Enhancing model predictive control by dynamic data reconciliation", AIChE, 48, 324-333, 2002.

Albuquerque, J.; and L.T. Biegler, "Data reconciliation and gross error detection for dynamic systems", AIChE, 42, 2841-2856, 1996.

Barbosa Jr, V.P.; M.R.M. Wolf; and R. Maciel Fo, "Development of data reconciliation for dynamic nonlinear system: application the polymerization reactor", Comp. Chem. Engng, 24, 501-506, 2000.

Binder, T.; L. Blank; W. Dahman; and W. Marquardt, "On the regularization of dynamic data reconciliation problems", J. Process Control, 12, 557-567, 2002.

Liebman, M.J.; T.F. Edgar; and L.S. Lasdon, "Efficient data reconciliation and estimation for dynamic process using nonlinear programming techniques", Comp. Chem. Engng, 16, 963-986, 1992.

Marlin, T.E., "Process control, design processes and control systems for dynamic performance", 2nd Ed., McGraw-Hill, 2000.

Narasimhan, S.; and C. Jordache, "Data reconcilation and gross error detection, an intelligent use of process data", Gulf publishing Company, Houston, 2000.

Oisiovici, R.M.; and S.L. Cruz, "State estimation of batch distillation column using an extended Kalman filter", Chem. Eng. Sci., 55, 4667-4680, 2000.

Ramamurthi, Y.; P.B. Sistu; and B.W. Bequette, "Control-relevant dynamic data reconciliation and parameter estimation", Comp. Chem. Engng, 17, 41-59, 1993.

Soderstrom, T.A.; T.F. Edgar; L.P. Russo; and R.E. Yong, "Industrial application of a large scale dynamic data reconciliation strategy", Ind. Eng. Chem. Res., 39, 1683-1693, 2000.

CHAPTER 5

Conclusions, Contributions and Recommendations

This thesis was devoted to the development of dynamic data reconciliation algorithms embedded within the structure of standard PID control loops. The controller performance was significantly enhanced by data reconciliation. The following areas were studied: the impact of measurement noise on the performance of controllers, the development of data reconciliation filters and their relevant applications, and of dynamic data reconciliation strategies and their implementations.

5.1 Conclusions and Contributions

Measurement noise may significantly deteriorate controller performance. The magnitude of the noise must be considered when tuning controllers using process models. Otherwise, the controllers are too aggressive, resulting in excessive variations in the controlled and manipulated variables. On the other hand, the detuning of controllers due to the presence of noise leads to more sluggish responses and lower controller performance. Controller performance can be partly regained by a data reconciliation algorithm embedded in the feedback loop, in order to obtain more consistent values of the controlled variables.

Data reconciliation uses process models as redundant information to compensate for measurement errors. Because process models are never perfect, model uncertainties

should be considered in the data reconciliation algorithm. The measurement and model errors partially compensate each other in the data reconciliation algorithm, so that more precise and consistent estimates of process variables are obtained.

In this work, two quasi-steady-state data reconciliation filters, SSDR and MWDR filters, were developed. To use quasi-steady-state data reconciliation filters, many process variables have to be measured to provide redundant information. The SSDR filter uses only the current set of measurements, whereas the MWDR filter also uses a series of past measurements to further attenuate the effect of noise. It was demonstrated that data reconciliation filters can significantly attenuate the effect of measurement noise, which allowed using more aggressive controllers and improving the overall performance of all control loops.

The dynamic data reconciliation (DDR) algorithm developed in this work can be implemented in real-time to enhance the controller performance. The dynamic data reconciliation provided more advantages than that of the quasi-steady-state data reconciliation filters, because it uses process dynamic models, rather than steady-state models. The dynamic models in the DDR algorithm can be phenomenological or empirical. For complex processes, empirical models can be developed by process identification techniques, and used in the DDR algorithm.

The two illustrative processes used in this work, the storage tank and the distillation column, revealed that the effectiveness and efficiency of the DDR algorithm depend on the accuracy of the models.

5.2 Recommendations

A number of recommendations can be made throughout this work.

1. The effects of the noise level on the performance of the data reconciliation algorithms embedded in the structure control loops should be investigated over a wide range of noise level in the future work.

2. The effectiveness of the data reconciliation algorithms and the controller performance were evaluated only for a disturbance of 20% step change in feed flow rate during this work. For other disturbances, for example, a step change in composition of the feed stream for the distillation column, or for setpoint changes of the controllers, the performance of the data reconciliation algorithms, along with the controllers, should be conducted in the future work.

3. It is advisable exploring and to using stochastic models in data reconciliation. For the dynamic data reconciliation using process input-output empirical models, it is difficult to estimate the variance of the models. It is suggested to first use large values of the variances, and then to tune these variances on-line until satisfactory performance of the DDR algorithm is obtained.

4. Empirical dynamic models were developed off-line for the distillation column, and used in the dynamic data reconciliation algorithm. The distillation column was assumed to be a linear-time-invariant (LTI) system. However, few plants are LTI systems. For example, the heat transfer coefficient for the reboiler of the distillation column may decrease due to fouling. In this case, the dynamic behavior predicted by the empirical models will deviate from the actual dynamics. As a result, the models will affect the performance of the DDR algorithm. On-line empirical model identification and dynamic data reconciliation performed simultaneously appear to be a promising method and should be investigated.

5. Because linear dynamic models are usually sufficient to capture the essence of process dynamics and they are easier to obtain than nonlinear models for most processes, Kalman filter could play an important role in the dynamic data reconciliation problems. Controller performance should be evaluated when Kalman filter is used. Kalman filter is a more complex algorithm than that of the DDR algorithm developed in this work. Evaluations of the DDR algorithm compared with the Kalman filter should be conducted.

6. Other newer methods, such as neural networks and fuzzy logic, are also attractive research areas when dealing with noisy data. Investigation of these two methods in data reconciliation are recommended.

APPENDIX A

Modeling of Distillation Dynamics

A.1 Models Used in the Simulation of the Distillation Column

Dynamic modeling and simulation of distillation processes have been widely studied during the past decades. Holland et al. (1983) introduced dynamic distillation models and their numerical methods, along with several illustrative examples. Gani et al. (1985) proposed generalized dynamic distillation models for both startup and continuous operation. Cameron et al. (1985) investigated numerical solutions for the dynamic distillation models. Choe et al. (1987) verified several assumptions in developing distillation models. The following assumptions have been generally made in modeling of distillation dynamics:

- Liquid holdups on trays, downcomers, reflux drum, and column base are considered to be well-mixed continuous stirred tanks.
- Tray efficiencies are often assumed to be 100%; otherwise, Murphree tray efficiencies are applied.
- The dynamics of the condenser, reboiler, measuring devices and control valves are negligible compared with the dynamics of the distillation column.

The binary (benzene-toluene) distillation column employed in this study is illustrated in Figure A.1. It has 21 sieve trays, a condenser, and a reboiler. The detailed geometric characteristics of the column are presented in Table A.1.

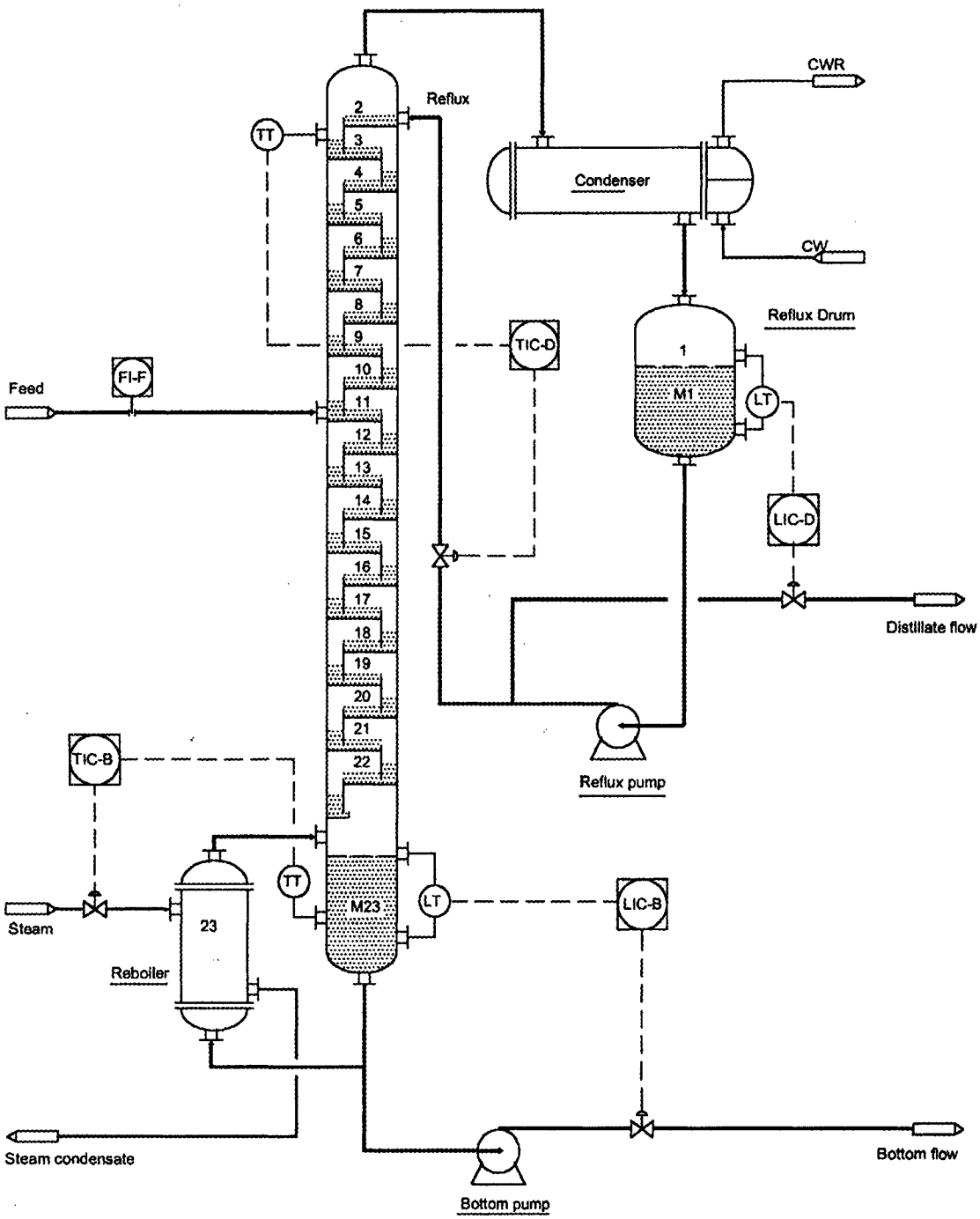

Figure A.1: Schematic diagram of benzene-toluene distillation column.

Table A.1 Critical dimensions of the benzene-toluene distillation column

Geometric Parameter	Unit	Size
Column diameter (D_t)	m	1.016
Tray spacing (S)	m	0.457
Downcomer area (A_D)	m^2	0.065
Tray net area (A_N)	m^2	0.743
Tray bubbling area (A_B)	m^2	0.678
Weir length (L_w)	m	0.681
Weir height (h_w)	m	0.051
Clearance under downcomer (h_{dl})	m	0.038
Downcomer width (W_{dc})	m	0.132
Hole diameter (D_H)	mm	4.8
Fractional hole area (A_f)	-	0.10
Hole pitch (Hp	mm	14.2
Tray thickness (W_t)	mm	3.4
Column base height (H_b)	m	1.270
Reflux drum diameter (D_r)	m	1.016
Reflux drum height (H_r)	m	1.016

In the application considered in this study, the benzene-toluene distillation column has one feed stream that is located at tray 11. The feed is a saturated liquid containing 0.3 mole fraction of benzene. The total feed flow rate is 100.0 kmol/h. The vapor at the top is totally condensed and collected in the reflux drum. One part of the liquid in the reflux drum is returned to the top of the column as reflux flow, while the other part is removed as the distillate flow. The distillation is operated approximately at atmospheric pressure. The distillate product stream is required to have a benzene concentration of 0.98 mole fraction, whereas the bottom product stream is required to have a benzene concentration of 0.01 mole fraction.

Kister (1990) has discussed several control configurations for a distillation column. The selection of a control scheme for a distillation column depends on the control objectives and the frequent disturbances to the column. In this control scheme, as shown in Figure A.1, two inferential feedback PI controllers, TIC-D and TIC-B, were used to control the top and bottom compositions of benzene by manipulating the reflux flow rate and the reboiler heat duty, respectively. Two other feedback PI controllers, LIC-D and LIC-B, were used to control the reflux drum and column base liquid levels by manipulating the distillate flow rate and the bottom product flow rate, respectively.

Since the benzene-toluene column was operated at a low pressure (atmospheric), the vapor holdup on trays can be neglected compared with the liquid holdup. The tray efficiency was assumed to be 100%. Based on the above assumptions, a distillation tray model for this column is presented in Figure A.2.

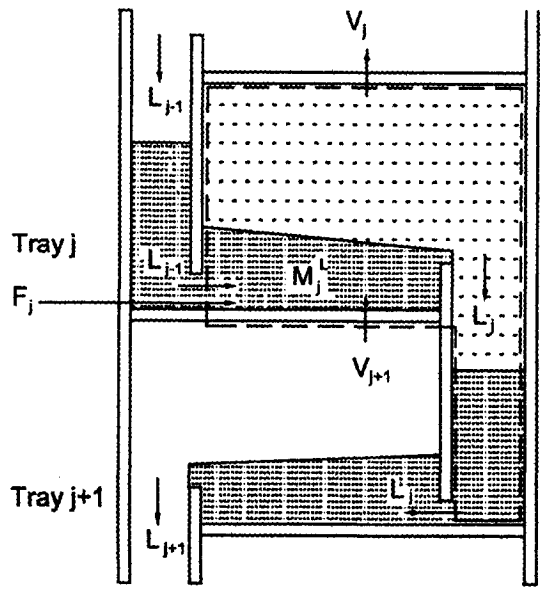

Figure A.2 Tray model for the benzene-toluene column

Mass and heat transfer occurring between the vapor and liquid phase takes place on each tray of the distillation column. The more volatile component, benzene, is found in a higher concentration in the vapor phase leaving each tray. The reverse is true for the less volatile component, toluene. A benzene-rich stream will be recovered at the top of the column whereas a toluene-rich stream will be recovered at the bottom of the column. In modeling all phenomena occurring on the distillation trays, a set of differential algebraic equations (DAEs) that describe the transient behavior of a distillation is defined. These equations include mass and heat balances, vapor-liquid equilibrium, tray hydraulics and physical properties.

The mass and heat balances are ordinary differential equations (ODEs); the other phenomenological equations such as the vapor-liquid equilibrium, tray hydraulics, and physical properties are algebraic equations.

Mass balance

As shown in Figure A.1, the condenser and reflux drum were considered as stage 1. The overall mass balance for stage 1 can be written as

$$\frac{dM_1}{dt} = V_2 - (L_1 + D) \tag{A.1}$$

where M_1 is the molar liquid holdup in the reflux drum, L_1 is the molar reflux flow, D is the molar distillate flow, V_2 is the molar vapor flow leaving stage 2. The component mass balance for benzene around stage 1 can be given as

$$\frac{d(M_1 x_{1,1})}{dt} = V_2 y_{1,2} - (L_1 + D)x_{1,1} \tag{A.2}$$

where $x_{1,1}$ is the mole fraction of benzene in the liquid of the reflux drum, $y_{1,2}$ is mole fraction of benzene in V_2.

The overall mass balance around stage j, $j \in [2,22]$, can be given by

$$\frac{d(M_j)}{dt} = V_{j+1} - V_j + L_{j-1} - L_j + F_j \tag{A.3}$$

where M_j is the liquid hold up on stage j, L_j is the liquid flow leaving stage j, V_j is the vapor flow leaving stage j, F_j is the feed flow to stage j. The component mass balance of benzene around stage j, $j \in [2,22]$, can be given by

$$\frac{d(M_j x_{1,j})}{dt} = V_{j+1} y_{1,j+1} - V_j y_{1,j} + L_{j-1} x_{1,j-1} - L_j x_{1,j} + F_j x_{F,j} \tag{A.4}$$

where $x_{1,j}$ and $y_{1,j}$ are liquid and vapor compositions of benzene leaving stage j, respectively. $x_{F,j}$ is the feed stream composition of benzene entering stage j.

For the reboiler and column base (j = 23), the overall mass balance can be written as

$$\frac{d(M_{23})}{dt} = L_{22} - V_{23} - B \tag{A.5}$$

where B is the bottom product flow. The component mass balance of benzene around the reboiler and column base is given by

$$\frac{d(M_{23} x_{1,23})}{dt} = L_{22} x_{1,22} - V_{23} y_{1,23} - B x_{1,23} \tag{A.6}$$

Heat balance

The heat balance around the condenser and reflux drum is given by

$$\frac{d(M_1 h_1)}{dt} = V_2 H_2 - (L_1 + D)h_1 - Q_C \tag{A.7}$$

where h_1 is the liquid specific molar enthalpy of L_1, H_2 is the vapor specific molar enthalpy of V_2, Q_C is the heat load of the condenser.

The heat balance around the stage j, $j \in [2, 22]$, can be written as

$$\frac{d(M_j h_j)}{dt} = V_{j+1} H_{j+1} - V_j H_j + L_{j-1} h_{j-1} - L_j h_j + F_j h_F \tag{A.8}$$

where h_j and H_j are the liquid and vapor specific molar enthalpies for L_j and V_j, respectively, h_f is the feed stream specific enthalpy. The liquid and vapor enthalpies are functions of temperature, pressure and composition. They need to be determined prior to solving equation (A.8). Equations used to calculate vapor and liquid enthalpies are given in Section A.2.

The heat balance around the reboiler and column base is given by

$$\frac{d(M_{23} h_{23})}{dt} = L_{22} h_{22} - V_{23} H_{23} - B h_{23} + Q_B \tag{A.9}$$

where Q_B is the heat load of the reboiler.

It should be noted that since the dynamics of the heat balance around a distillation stage are relatively fast compared with the dynamics of the compositions, the heat balances have often been assumed to be at pseudo steady-state in most literature and the derivatives in equation (A.8) are set to be zeroes. Ranzi et al. (1988) showed that neglecting the time dynamics in the heat balance caused large discrepancies with results

using the rigorous differential balance equations. The differential model equations were used in this simulation.

Vapor-liquid phase equilibrium

The vapor and liquid phases on each stage *j* are assumed to be in equilibrium. The relationship between the vapor and liquid compositions on stage *j* is described by the following equation

$$y_{i,j} = K_{i,j} x_{i,j} \tag{A.10}$$

where $K_{i,j}$ is the vapor-liquid equilibrium constant for component *i*. The values of the equilibrium constants are complex functions of system temperature and pressure. The modeling of K-values is presented in Section A.3.

Tray hydraulics

The liquid flow rate leaving tray *j* was correlated with the height of liquid over the weir. The liquid height over the weir depends on the liquid holdup on a tray, physical properties of the vapor and liquid and tray geometry. In general

$$L_j = L(\rho_j^V, \rho_j^L, \eta_j^V, \eta_j^L, \sigma_j^L, \mathbf{G}) \tag{A.11}$$

where L_j is the liquid flow leaving tray *j*, ρ_j^V and ρ_j^L are the vapor and liquid densities, respectively, η_j^V and η_j^L are the vapor and liquid viscosities, respectively, σ_j^L is the liquid surface tension, and $\mathbf{G}$ is the vector of tray geometry defined in Table A.1.

For the vapor to flow through a tray, a pressure difference across the tray must exist. The pressure drop across each tray, given the vapor flow rate, can be calculated and allows to determined pressure profile along the column. The pressure drop of tray j is a function of the vapor flow, physical properties of the vapor and liquid, and the tray geometry

$$\Delta P_j = P(V_{j+1}, \rho_j^V, \rho_j^L, \eta_j^V, \eta_j^L, \sigma_j^L, \mathbf{G}) \tag{A.12}$$

With the pressure drop calculated for each tray, the column pressure profile can be calculated by the simple equation

$$P_j = P_{j-1} + \Delta P_{j-1} \tag{A.13}$$

Details for modeling tray hydraulics are presented in Section A.4.

Physical properties

Physical properties of the vapor and the liquid on tray j need to be calculated. These physical properties are complex correlations of system temperature, T_j, pressure, P_j, and compositions, $x_{i,j}$ and $y_{i,j}$. For simplicity, these equations are expressed in the functional forms:

$$\rho_j^V = \rho^V(T_j, P_j, y_{i,j}) \tag{A.14}$$

$$\rho_j^L = \rho^L(T_j, P_j, x_{i,j}) \tag{A.15}$$

$$\eta_j^V = \eta^V(T_j, P_j, y_{i,j}) \tag{A.16}$$

$$\eta_j^L = \eta^L(T_j, P_j, x_{i,j}) \tag{A.17}$$

$$\sigma_j^L = \sigma^L(T_j, P_j, x_{i,j}) \tag{A.18}$$

The development of these physical property models for vapor and liquid mixtures are described in Section A.2.

A.2 Modeling of Physical Properties

Mathematical models of physical properties developed in the literature have some theoretical basis, but heavily rely on empiricism. In modeling physical properties of mixtures, basic physical properties of pure components such as molecular weight, critical values of temperature, pressure, and other coefficients of empirical correlations, are usually required. For the two components, benzene and toluene, the basic physical properties are listed in Appendix D.

Density of Vapor and Liquid

Since the column is operated at low pressure (atmospheric), the vapor is assumed to obey the ideal gas law. Therefore, the ideal gas state equation was used to calculate the vapor density

$$\rho^V = \frac{P}{RT}\sum_{i=1}^{2} M_i y_i \tag{A.19}$$

where P and T are the system pressure and temperature, respectively. R is the ideal gas constant, y_i is the mole fraction of component i in the vapor, and M_i is the molecular weight of component i.

If an accurate value of the pure liquid density is available at a given temperature, the density of pure liquid can be calculated by (Reid and Sherwood, 1966)

$$\rho_0^L = 2\rho_C + \frac{T_C - T}{T_C - T_R}(\rho_R - 2\rho_C) - \rho_0^V \tag{A.20}$$

where ρ_C is the critical density, T_C is the critical temperature, ρ_R is a reference liquid density at the reference temperature T_R , and ρ_0^V is the density of the pure vapor at the system temperature and pressure.

The liquid mixture of benzene-toluene was treated as an ideal solution. Its density can be calculated by the linear mixing rule

$$\rho^L = \sum_{i=1}^{2} \rho_{0,i}^L x_i \tag{A.21}$$

where $\rho_{0,i}^L$ is the density of pure liquid of component i, x_i is the mole fraction of component i in the liquid.

Viscosity of Vapor and Liquid

Vapor viscosity is a function of system temperature and pressure. However, at low pressures the effect of pressure can be neglected. Yaws (1995) gives correlations of vapor viscosity of pure components as a function of temperature

$$\eta_0^V = a + bT + cT^2 \tag{A.22}$$

where a, b, and c are coefficients given in Appendix D and T is the system temperature.

The viscosity of a vapor mixture is usually not linear in compositions. Reid and Sherwood (1966) recommended an equation for the viscosity of the binary vapor mixture

$$\eta^V = \frac{\eta_{0,1}^V}{1+(y_2/y_1)\Phi_{12}} + \frac{\eta_{0,2}^V}{1+(y_1/y_2)\Phi_{21}} \tag{A.23}$$

where $\eta_{0,1}^V$ and $\eta_{0,2}^V$ are the vapor viscosity of pure benzene and toluene, respectively. y_1 and y_2 are the mole fractions of benzene and toluene in the vapor, respectively. The Φ_{12} and Φ_{21} are calculated by

$$\Phi_{12} = \frac{\left[1+(\eta_{0,1}^V/\eta_{0,2}^V)^{0.5}(M_2/M_1)^{0.25}\right]^2}{\sqrt{8}\left[1+(M_1/M_2)\right]^{0.5}} \tag{A.24}$$

$$\Phi_{21} = \frac{\left[1+(\eta_{0,2}^V/\eta_{0,1}^V)^{0.5}(M_1/M_2)^{0.25}\right]^2}{\sqrt{8}\left[1+(M_2/M_1)\right]^{0.5}} \tag{A.25}$$

where M_1 and M_2 are the molecular weights of benzene and toluene, respectively.

At low pressures, the effect of pressure on liquid viscosity can be neglected. The Yaws' (1995) empirical correlations for pure liquid viscosity were used

$$Log_{10}\eta_0^L = a + \frac{b}{T} + cT + dT^2 \tag{A.26}$$

where a, b, c, and d are coefficients of the polynomial given in Appendix D.

The viscosity of a liquid mixture is not linear in compositions. For this binary liquid mixture, the viscosity was calculated by Reid and Sherwood (1966):

$$\eta^L = (x_1\sqrt{\eta_{0,1}^L} + x_2\sqrt{\eta_{0,2}^L})^2 \tag{A.27}$$

Liquid Mixture Surface Tension

The surface tension of pure liquids is a function of system temperature and pressure. Neglecting the effect of pressure on surface tension, the surface tension of pure liquids was calculated by Reid and Sherwood (1966):

$$\sigma_0^L = kT_C^{1/3} P_C^{2/3} (1 - T_r)^n \tag{A.28}$$

where T_C and P_C are the critical temperature and pressure, respectively, and $T_r (= T / T_C)$ is the reduced temperature. Parameters k and n are specific constants for each compound. For benzene, $k = 2.9997$ and $n = 1.2291$. For toluene, $k = 3.2931$ and $n = 1.3220$

The liquid surface tension of the benzene-toluene mixture was considered to be linear in composition:

$$\sigma^L = \sum_{i=1}^{2} x_i \sigma_{0,i}^L \tag{A.29}$$

Enthalpy of Vapor and Liquid

The pure liquid enthalpy was calculated by the integral formula

$$h_0^L = h_{0,ref}^L + \int_{T_{ref}}^{T} Cp_L dT \tag{A.30}$$

where h_0^L is the pure liquid enthalpy at T temperature, $h_{0,ref}^L$ is the reference enthalpy at temperature T_{ref} (298.15 K). Cp_L is the liquid heat capacity that is expressed in the polynomial form (Felder, 2000), $Cp_L = a + bT$, for the benzene and toluene. The coefficients *a* and *b* are given in Appendix D.

For liquid mixtures having *n* components, the liquid enthalpy is calculated by

$$h^L = \sum_{i=1}^{n} x_i h_{0,i}^L + \Delta h_{soln} \tag{A.31}$$

where x_i is the mole fraction of component i in the liquid mixture, $\Delta h_{so\ln}$ is the heat of solution of the liquid mixture. For the ideal liquid solution of benzene-toluene, the $\Delta h_{so\ln}$ was neglected.

The pure vapor enthalpy was calculated using the following integral formula

$$H_0^V = H_{0,ref}^V + \int_{T_{ref}}^{T} Cp_V dT \tag{A.32}$$

where H_0^V is the pure vapor enthalpy at temperature T, $H_{0,ref}^V$ is the reference enthalpy at temperature T_{ref} (298.15 K). Cp_V is the vapor heat capacity that is given by the polynomial form (Felder, 2000), $Cp_V = a + bT + cT^2 + dT^3$, for the benzene and toluene. The coefficients *a, b, c, d,* are given in Appendix D.

For vapor mixtures having n components, the vapor enthalpy can be calculated by

$$H^V = \sum_{i=1}^{n} H_{0,i}^V y_i + \Delta H_{so\ln} \tag{A.33}$$

where y_i is the mole fraction of component i in the vapor, $\Delta H_{so\ln}$ is the heat of the solution of vapor mixture that was neglected for the hydrocarbon mixtures.

A.3 Modeling of Vapor-Liquid Equilibrium

The vapor-liquid system of benzene-toluene was treated as an ideal binary system, where Dalton's law and Raoult's law can be applied. Applying Dalton's law gives

$$p_i = y_i P \tag{A.34}$$

where p_i is the partial pressure of component i in the vapor, y_i is the mole fraction of the component i in the vapor, and P is the system pressure. Applying Raoult's law gives

$$p_i = x_i p_i^0 \tag{A.35}$$

where p_i^0 is the vapor pressure of pure component i at system temperature. Combining equations (A.34) and (A.35) yields the vapor-liquid equilibrium constant

$$K_i = \frac{y_i}{x_i} = p_i^0 / P \tag{A.36}$$

The vapor pressure of pure component i, p_i^0, is a function of system temperature, and it was calculated by the Antoine equation (Boublik, 1984):

$$Log_{10} p_i^0 = A_i - \frac{B_i}{T + C_i} \tag{A.37}$$

where A_i, B_i, and C_i are the Antoine constants of pure component i that are given in Appendix D.

A.4 Modeling of Tray Hydraulics

Tray hydraulic models are complex functions of vapor and liquid flow, physical properties, and tray geometry. There is no theoretical equation available that gives a general model describing tray hydraulics. Empirical correlations and charts have been developed in the literature to describe the behavior of tray hydraulics.

In order to develop the necessary empirical correlations, it is necessary to be familiar with tray geometry. Figure A.3 illustrates a typical diagram of a sieve tray. The bubbling area, A_B, is the active area of the tray available for contacting vapor and liquid. The

downcomer area, A_D, is the area available for liquid flowing to the next tray. A_{da} is the area under the downcomer where the tray receives the liquid from the tray above.

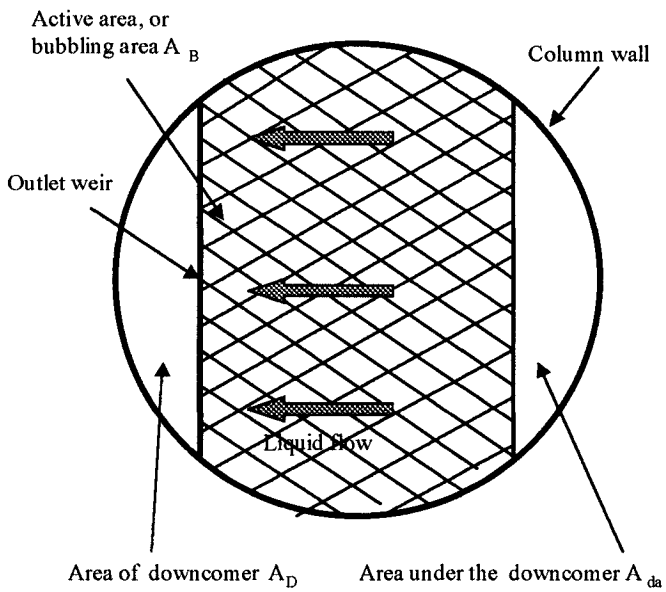

Figure A.3 A typical tray geometry

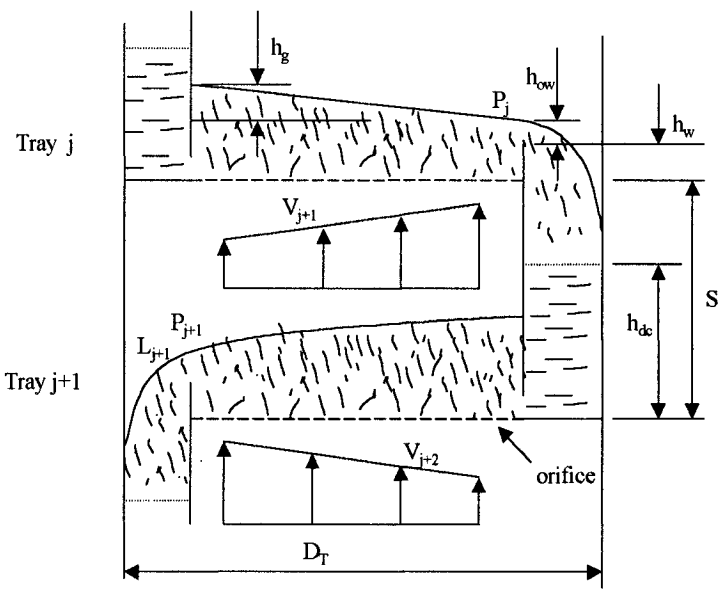

Figure A.4 Hydrodynamics of vapor and liquid on trays

Tray Pressure Drop

Figure A.4 shows a typical hydrodynamic sketch of vapor and liquid flow around a tray. When the vapor goes up through a tray, it has to overcome the resistances exerted by the tray itself and the liquid layer. The total pressure drop across a tray is the sum of the pressure drop across the tray orifices and the pressure drop through the aerated liquid

$$h_t = h_d + h_l \tag{A.38}$$

where h_t is the total tray pressure drop, h_d is the dry pressure drop resulting from the vapor passing through tray orifices, h_l is the pressure drop of vapor passing through the aerated liquid on the tray. The dry pressure drop h_d is given by (Kister, 1992)

$$h_d = 1.86(\frac{U_h}{C_v})^2 \frac{\rho_G}{\rho_L} \tag{A.39}$$

where, U_h is the vapor hole velocity, based on total orifice area on a tray, C_v is the orifice coefficient obtained by the chart shown in Figure A.5. ρ_G and ρ_L are the vapor and liquid densities respectively.

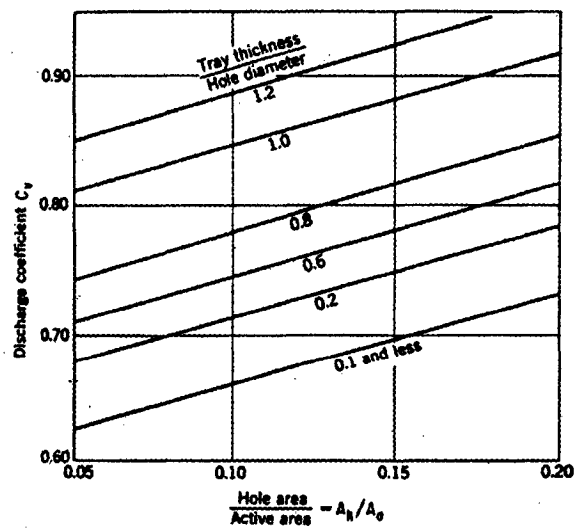

Figure A.5 Orifice coefficient chart (Kister, 1992)

To calculate the pressure drop through the aerated liquid, h_l, the Fair's pressure drop correlation was used (Kister, 1992)

$$h_l = \beta h_c \tag{A.40}$$

where β is the aeration factor, dimensionless, and h_c is the liquid height on the tray. The aeration factor β is given by a chart shown in Figure A.6 where it is indicated that factor β is a function of F-factor, $F_{ga} = U_a \rho_G^{0.5}$, in which U_a is the vapor superficial velocity based on bubbling area, and ρ_G is the vapor density.

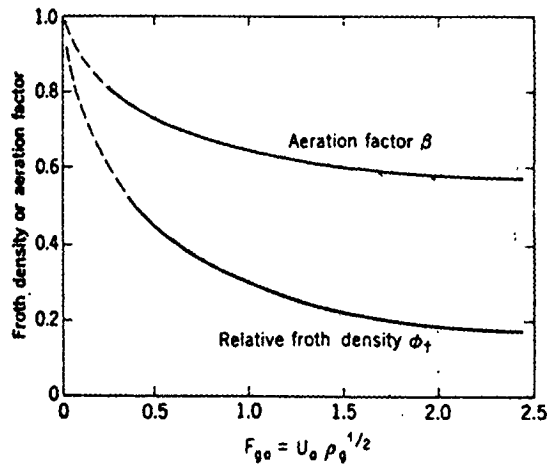

Figure A.6 Tray aeration factor prediction chart (Kister, 1992)

The linear least-squares technique was used to give a mathematical equation by fitting sets of sampled data from the chart in Figure A.6. A third order polynomial equation was used. The fitted model is

$$\beta = 0.995 - 0.6293 U_a \rho_G^{0.5} + 0.3314 (U_a \rho_G^{0.5})^2 - 0.06 (U_a \rho_G^{0.5})^2 \tag{A.41}$$

The liquid height on a tray, h_c, is composed of three layers, the height of weir, h_w, the height of liquid above the weir, h_{ow}, and the liquid hydraulic gradient, h_g

$$h_c = h_w + h_{ow} + h_g / 2 \tag{A.42}$$

The liquid height over the outlet weir, h_{ow}, was calculated by the Francis weir formula (Kister, 1992)

$$h_{ow} = 0.48 F_w (\frac{Q_{Liq}}{L_w})^{2/3} \tag{A.43}$$

where Q_{Liq} is the liquid load of the tray, L_w is the length of outlet weir. F_w is the weir construction factor that is a function of $Q_{liq} / L_w^{2.5}$ obtained by a chart shown in Figure A.7. In simulation, it is necessary to have a mathematical model for $F_w \sim Q_{liq}/L_w^{2.5}$.

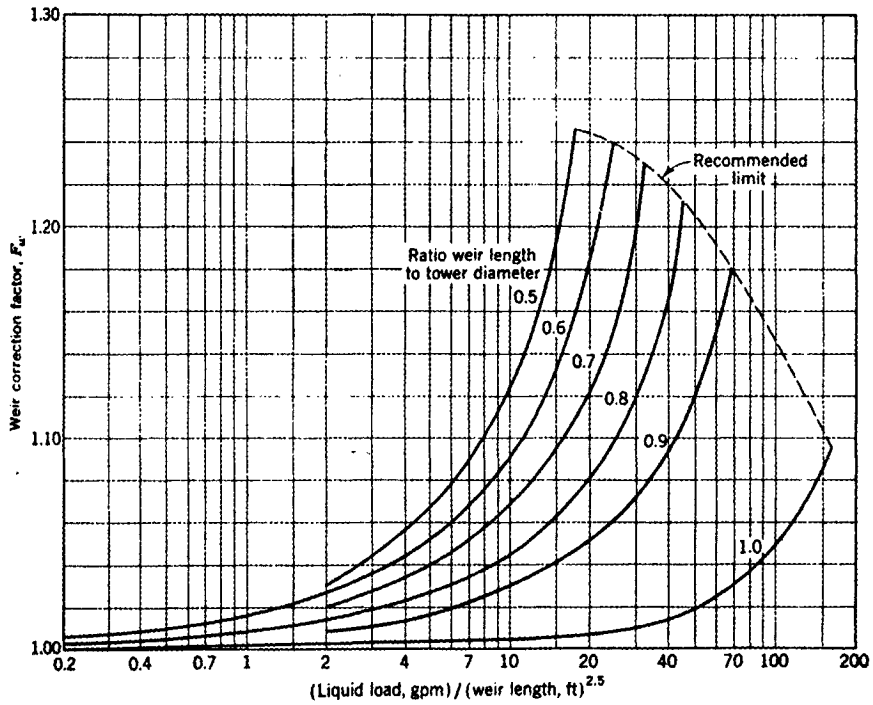

Figure A.7 Weir correction factor chart (Kister, 1992).

Using least-squares technique, a third-order polynomial equation was used to fit sampled data from the chart, the model equation obtained is

$$F_w = 1.0071 + 0.0076(Q_{Liq.}/L_w)^{2.5} - 0.0002[(Q_{Liq.}/L_w)^{2.5}]^2 + 6.0\times10^{-6}[(Q_{Liq.}/L_w)^{2.5}]^3 \quad \text{(A.44)}$$

The hydraulic gradient, h_g, is the head of liquid necessary to overcome the fractional forces in the path of liquid on a tray. The hydraulic gradient can be negligible for most sieve trays, but it cannot be neglected for large diameter trays whose hydraulic gradient has a significant effect on pressure drop calculations. The empirical correlation for the hydraulic gradient, h_g, is given as (Kister, 1992)

$$h_g = \frac{12 f U_f^2 L_f}{g R_h} \quad \text{(A.45)}$$

where, f is the frictional factor (dimensionless), U_f is the velocity of the aerated mass across the tray, L_f is the length of flow path across the tray, g is the acceleration factor due to gravity, and R_h is the hydraulic radius of the aerated mass.

In equation (A.45) the f, U_f, and R_h are unknown variables and need to be calculated. The U_f is given as (Kister, 1992)

$$U_f = \frac{1}{37.4}\frac{Q_L}{h_l}\frac{L_w}{D_f} \quad \text{(A.46)}$$

where $Q_L = Q_{Liq}/L_w$, D_f is the arithmetic average between tower diameter and weir length, h_l is the liquid height on the tray. Since h_l is calculated from h_g, a trial-and-error calculation procedure is required to solve the problem. To simplify the problem, h_l can be approximated as $h_l = 0.65(h_w + h_{ow})$ in calculation of U_f.

The hydraulic radius of the aerated mass R_h is given as (Kister, 1992)

$$R_h = \frac{h_f D_f}{2h_f + 12D_f} \tag{A.47}$$

where h_f is the froth height on the tray estimated by $h_f = 2.6(h_w + h_{ow})$.

The friction factor f in equation (A.45) is a function of Reynolds number $N_{\mathrm{Re},h} = (R_h U_f \rho_L)/\eta_L$ where η_L is the liquid viscosity. For this column, the model equation of f is given by

$$Log_{10}(f) = 3.6422 - 1.0568 Log_{10}(N_{\mathrm{Re},h}) \tag{A.48}$$

Height of Liquid in the Downcomer

As shown in Figure A.4, the liquid is backed up into the downcomer by tray pressure drop, liquid height on the tray, and frictional losses in the downcomer apron. Making a pressure balance around tray *j* gives

$$h_{dc,j} = h_{cliq,j+1} + h_{t,j} + h_{da,j} \tag{A.49}$$

where $h_{cliq,j+1}$ is the clear liquid height on tray *j+1* calculated by (A.40), $h_{t,j}$ is the tray pressure drop calculated by equation (A.38), $h_{da,j}$ is the head loss due to liquid flow under the downcomer apron, calculated by the correlation (Kister, 1992)

$$h_{da,j} = 0.03\left(\frac{G_{Liq}}{100A_{da}}\right)^2 \tag{A.50}$$

where G_{Liq} is the liquid flow rates on the tray *j*, A_{da} is the area under downcomer apron.

Liquid flow

As shown in Figure A.4, the liquid volume on a tray is composed of two parts: the tray bubbling area and in the downcomer. The total volume of clear liquid on a tray can be written as

$$V = A_b\beta(h_w + h_{ow} + 0.5h_g) + A_d h_{dc} \tag{A.51}$$

where V is the volume of clear liquid on the tray, A_b is the bubbling area on the tray, A_d is the downcomer area, β is the aeration factor, h_w is the height of weir, h_{ow} is the height of liquid over outlet weir, h_g is the hydraulic gradient, h_{dc} is the height of clear liquid in downcomer. Rearranging equation (A.51) gives

$$h_{ow} = \frac{V - A_d h_{dc}}{A_b\beta} - h_w - 0.5h_g \tag{A.52}$$

Putting $V = \frac{M}{\rho_L}$ where M is the liquid molar holdup on the tray, ρ_L is the molar density of the clear liquid, into equation (A.52) and applying the modified Francis weir formula $h_{ow} = 0.01F_w\left(\frac{Q}{L_w}\right)^{2/3}$ yields

$$0.01F_w\left(\frac{Q}{L_w}\right)^{2/3} = \frac{M/\rho_L - A_d h_{dc}}{A_b\beta} - h_w - 0.5h_g \tag{A.53}$$

Rearranging (A.53) yields

$$Q = \left[\frac{(M/\rho_L - A_d h_{dc})/A_b\beta - h_w + 0.5h_g)}{0.01F_w}\right]^{1.5} L_w \tag{A.54}$$

It should be noted that in equation (A.54), the hydraulic gradient h_g, the head of clear liquid in downcomer h_{dc}, and the weir correlation factor F_w are functions of liquid load Q. Therefore, equation (A.54) is an implicit form in Q. Note that in (A.54) Q_{Liq} is in m^3/h, and L_w is in m.

A.5 Numerical Solutions to Distillation Dynamic Models

The dynamic models of the distillation column as described above consist of coupled differential-algebraic equations (DAEs) with the characteristics:

- Ordinary differential equations (ODEs) originate from mass and heat balances.
- Highly nonlinear algebraic equations (AEs) come from phenomenological quantities, such as the physical properties and tray hydraulics.
- There are far more AEs than ODEs.
- The DAEs are of explicit or implicit types.

The representation of the DAEs can be written in the general form

$$\mathbf{f}(\frac{d\mathbf{y}}{dt}, \mathbf{y}, \mathbf{x}, t) = \mathbf{0} \tag{A.55}$$

$$\mathbf{g}(\mathbf{y}, \mathbf{x}, t) = \mathbf{0} \tag{A.56}$$

where $\mathbf{y}$ is the vector of differential variables, $\mathbf{x}$ is the vector of algebraic variables. Two approaches have been proposed in the literature to solve the DAEs. One approach is the *ODEs method*, and the other is the *simultaneous solution method* (Pantelides and Barton, 1992).

The ODE method considers the coupled DAEs consisting of ODEs with boundary conditions or constraints imposed by equation (A.56). In the integration of the ODEs, the algebraic equations are first solved to obtain the algebraic variables in terms of the differential ones. Then, the algebraic variables are substituted into the ODEs in to convert the DAEs into a set of ODEs in the differential variables, $\mathbf{y}$. The reduced form of the ODEs can be subsequently solved by conventional ODE algorithms. The simultaneous solution method was evolved from the numerical integration of the ODEs. One algorithm for the simultaneous solution method is the BDF method that is based on Gear's backward differentiation formula (Marquardt, 1995). In this work, an explicit Euler's method (Luyben, 1990) was used to integrate the ODEs with the ODEs method.

References

Boublik, T.; V. Fried; and E. Hala, "The vapor pressure of pure substances", Elsevier, 1984.

Cameron, I.T.; C.A. Ruiz; and R. Gani, "A generalized model for distillation columns – I: numerical and computational aspects", Comp. Chem. Engng, 10, 199-211, 1986.

Choe, Y.; and W.L. Luyben, "Rigorous dynamic models of distillation columns", Ind. Eng. Chem. Res., 26, 2158-2161, 1987.

Felder, R.; and R.W. Rousseau, "Elementary principles of chemical processes", third edition, John Wiley & Sons, 2000.

Gani, R.; C.A. Ruiz; and I.T. Cameron, "A generalized model for distillation columns – I: model description and applications", Comp. Chem. Engng, 10, 181-198, 1986.

Holland, C.D.; and A.I. Liapis, "Computer methods for solving dynamic separation problems, McGraw-Hill, 1983.

Kister, H.Z., "Distillation operation", McGraw-Hill, 1990.

Kister, H.Z., "Distillation design", McGraw-Hill, 1992.

Lide, D.R., "CRC handbook of chemistry and physics", 81st Ed., CRC Press, 2000-2001.

Luyben, W.L., "Process modeling, simulation, and control for chemical engineers", 2nd Ed., McGraw-Hill, 1990.

Marquardt, W., "Numerical methods for the simulation of differential-algebraic process models", NATO ASI series E: Applied sciences, Methods of Model Based Process Control, vol. 293, pp.41-79, 1995.

Pantelides, C.C.; and P.I. Barton, "Equation-oriented dynamic simulation current status and future perspectives", Suppl. to Comp. Chem. Engng, European Symposium on Computer Aided Process Engineering-2, October 1992.

Reid, R.C.; and T.K. Sherwood, "The properties of gases and liquid", 2nd Ed., McGraw-Hill, 1966.

Ranzi, E.; M. Rovaglio; T. Faravelli; and G. Biardi, "Role of energy balances in dynamic simulation of multicomponent distillation columns", Comp. Chem. Engng, 12, 783-786,1988.

Yaws, C.L., "Handbook of transport properties data", Gulf publishing, 1995.

APPENDIX B

Validation of Distillation Dynamic Simulator

B.1 Simulation of Open-loop Responses

In order to validate the dynamic distillation simulator, the open-loop responses of the main process variables were first simulated. To perform open-loop simulations, the column was allowed to reach steady state. Then, without control action, a step change in each individual input variable (distillate flow, reflux flow, bottom flow, reboiler heat duty, feed flow and feed composition) was introduced to the column, and the responses of the four controlled variables (reflux drum level, top temperature, base level and bottom temperature) were recorded as a function of time. It should be noted that the step changes were 20% of the steady-state values, except a 10% step change in reboiler heat duty was used to generate the response of the top temperature. The open-loop simulation results are presented in Figure B.1.

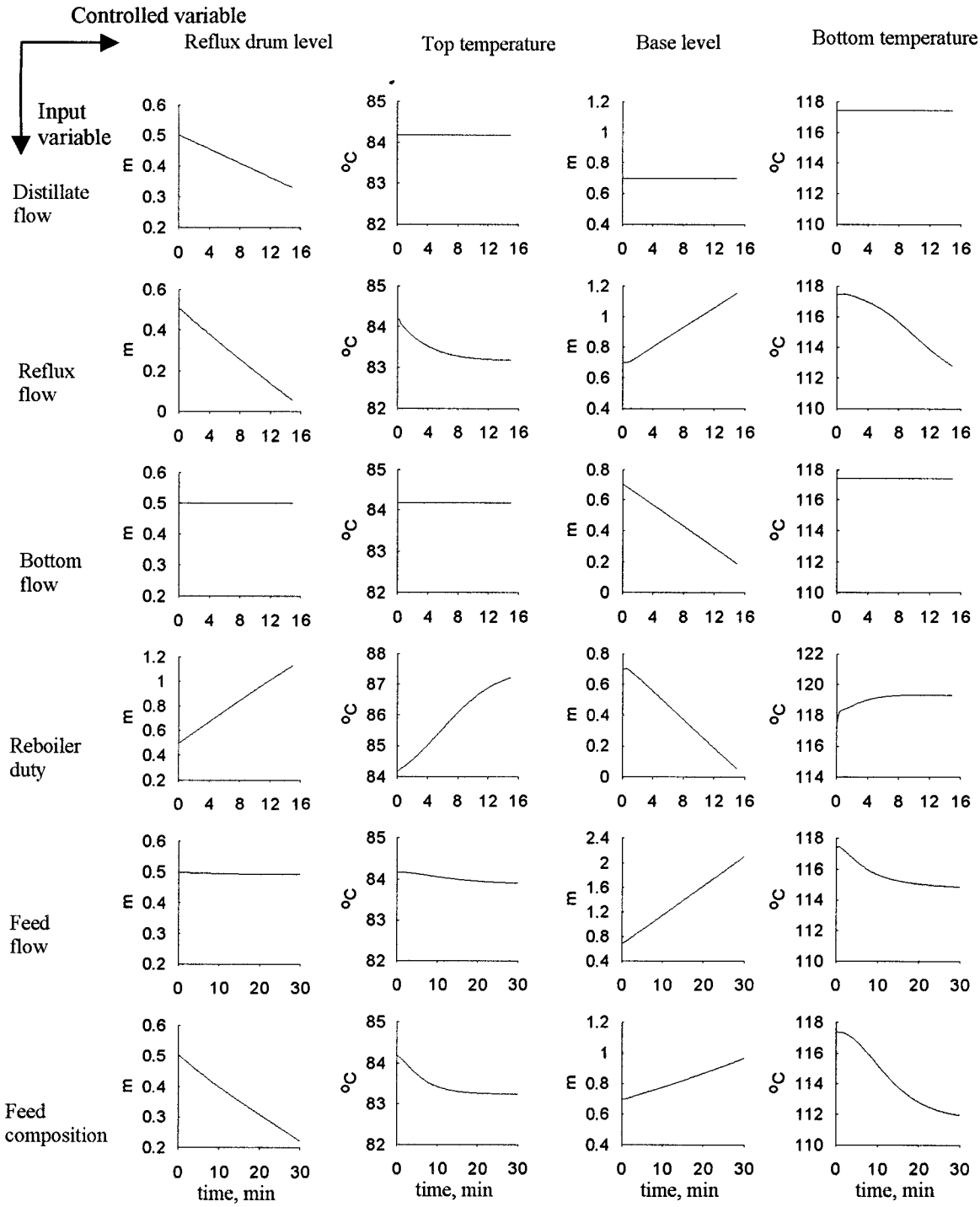

Figure B.1 Simulation results of open-loop responses of the distillation column.

Responses to a step change in distillate flow

With the step change of the distillate flow, the liquid level of the reflux drum decreased monotonously as a ramp, whereas, it had no impact on the other three controlled variables. Indeed, the response curves of the top temperature, base level and bottom temperature are horizontal lines (see the first row in Figure B.1).

Responses to a step change in reflux flow

With a step change in reflux flow rate, the reflux drum liquid level decreased as a ramp, and the top temperature decreased as a first-order system due to a larger amount of colder liquid returned to the top of the column. The base level increased, but with some time delay because the liquid has to pass throughout the column tray by tray from top to bottom before reaching the base. The bottom temperature decreased as a first-order-with-dead-time system, but it can also be modeled as a pure-integrator-with-dead time system (see the second row in Figure B.1).

Responses to a step change in bottom flow

The step change in bottom flow had no impact on the reflux drum liquid level, the top temperature, and the bottom temperature. The only affected controlled variable was the base level that decreased as a ramp (see the third row in Figure B.1).

Responses to a step change in reboiler heat duty

With a step increment of reboiler heat duty, the reflux drum level increased. The reason is that more fluid was vaporized by supplying more heat to the reboiler. As a result, more

vapor condensed at the top in the condenser, and more liquid flow entered into the reflux drum. The top temperature increased as a pure-integrator-with-dead-time, or a first-order-with-dead-time system. This resulted from the vapor fluid conveying more toluene to the top. The column base level remained constant for a short time, then it decreased as a ramp. The reason of the decrease of the base level is that there was more fluid in the base being vaporized following the increased reboiler heat duty. The bottom temperature was increased with the increase of reboiler heat duty. The increase of bottom temperature was fast at the beginning, then changed slowly until reaching the new steady state (see the fourth row in Figure B.1).

Responses to a step change in feed flow

The reflux drum level remained constant with the increase of feed flow, while the column base level increased monotonously with the same external disturbance because the feed is a saturated liquid. The top temperature displayed a slight decrease because the increment in the feed flow leads to increase the benzene concentration at the base of the column so that the vapor flowing up from the column base contained more volatile benzene. The bottom temperature decreased as a first-order-with-dead-time system. The decrease of the bottom temperature is due to the increase of the benzene concentration at the column base (see the fifth row in Figure B.1).

Responses to a step change in feed composition

With a step change in benzene mole fraction in the feed stream, the level of reflux drum decreased, whereas the level of the column based was increased. Both the top and bottom

temperatures decreased as first-order-with-dead-time systems. The decrease of the bottom temperature resulted from an increase in benzene concentration at the column base. Because the heat of vaporization of the liquid mixture increases with the decrease of the liquid temperature, the vapor flow leaving the column base decreased by given the same amount of heat to the reboiler. The reduction in the vapor flow rate led to an increase of the level at the base of the column and a decrease in the level of the reflux drum. The decrease of the top temperature resulted from the phenomenon that a larger quantity of benzene was brought to the top of the column by the vapor containing more benzene (see the sixth row in Figure B.1).

B.2 Simulation of Closed-loop Responses

As the simulation results of the open-loop responses (Figure B.1) indicated, the dynamics of the two liquid levels are characterized by pure integrator processes, whereas the dynamics of the two temperatures are characterized by first-order-with-dead-time systems. Closed-loop simulations were performed using a sampling interval of 30 seconds. The Ziegler-Nichols tuning rules (Ogunnaike et al., 1994) were used to give the initial values of the four PI controller parameters. The values of these parameters are shown in Table B.1.

Table B.1 Initial controller tuning parameters

Controller	LIC-D		TIC-D		LIC-B		TIC-B	
	K_C $kmol.h^{-1}.m^{-1}$	τ_I s	K_C $kmol.h^{-1}.°C^{-1}$	τ_I s	K_C $kmol.h^{-1}.m^{-1}$	τ_I s	K_C $MJ.h^{-1}.°C^{-1}$	τ_I s
Parameter values	-960.0	93.0	-87.0	96.6	-960.0	93.0	850.7	143.2

To obtain the close-loop responses, the simulator was first allowed to reach steady state at the nominal operation point before introducing an increase of 20% in feed flow rate. The close-loop responses of the four controlled and manipulated variables are presented in Figure B.2.

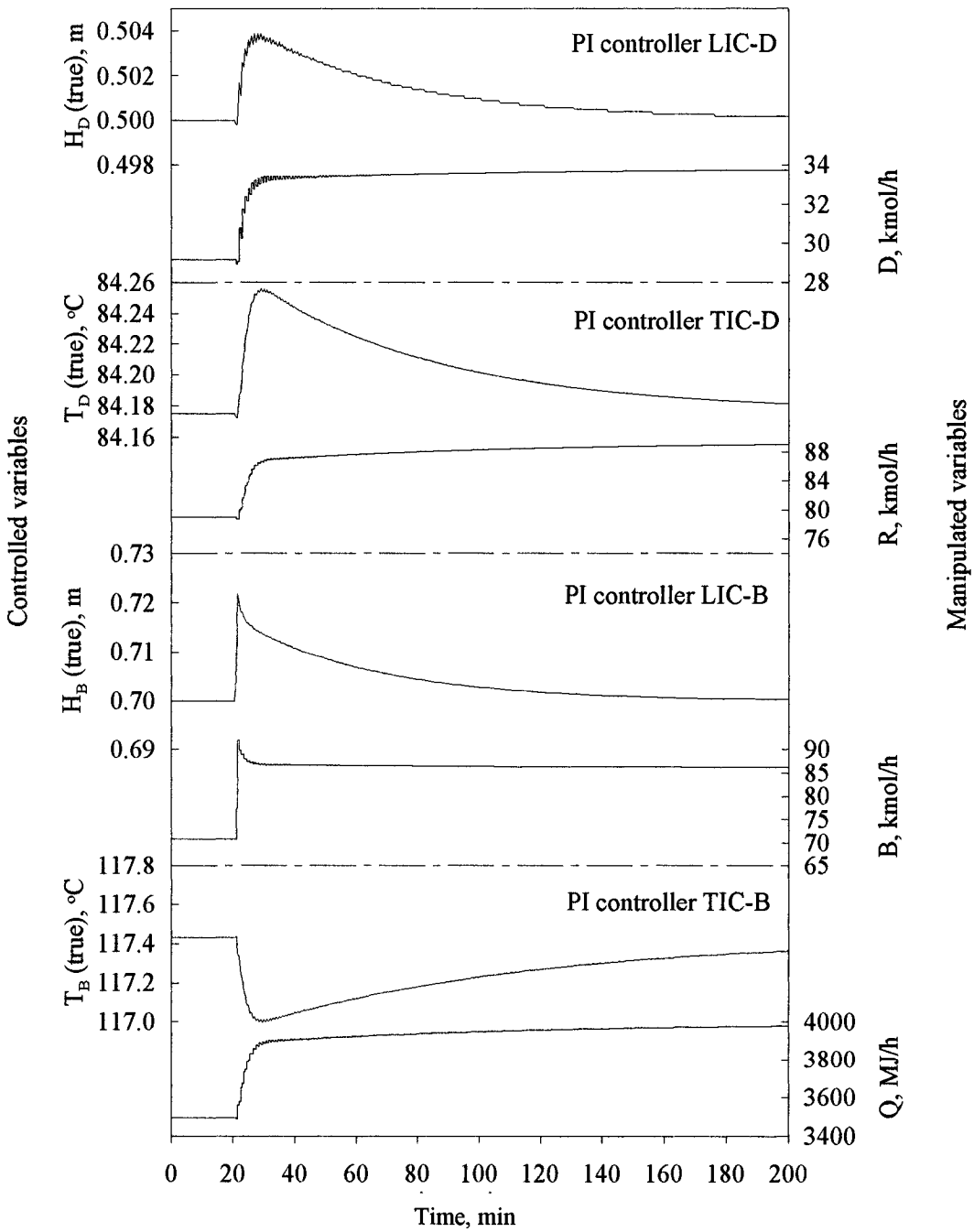

Figure B.2 Simulation results of closed-loop responses for a 20% step change in feed flow.

As shown in Figure B.2, an increase in feed flow led to an increase in the level of the base of the column. Because the liquid level exceeded its setpoint, the controller (LIC-B) responded by increasing the bottom flow rate (the manipulated variable) to maintain the level at its setpoint. It took approximately 120 minutes for the base level to return to its steady-state value after the disturbance (see curves of H_B and B in Figure B.2).

The temperature of the column base decreased with an increase in the feed flow rate. As a result, the controller (TIC-B) increased the heat supply to the reboiler (the manipulated variable) to return the bottom temperature to its setpoint. The response of the bottom temperature was relatively sluggish, as it took approximately 200 minutes for the bottom temperature to reach its steady state value (see curves of T_B and Q in Figure B.2).

The vapor flow inside the column increased as more heat was supplied to the reboiler during the transient of the column dynamics. At the distillate end of the column, the increment of the vapor flow entrained more toluene to reach to the top of the column. Therefore, the temperature at the top of the column increased due to higher toluene concentration on the top trays. The controller (TIC-D) consequently increased the reflux flow (the manipulated variable) to the column to control the top temperature at its setpoint. It took about 200 minutes for the top temperature to reach its steady state value after the feed flow perturbation (see curves of T_D and R in Figure B.2).

The reflux drum level increased because the vapor flow rate increased. The controller (LIC-D) increased the distillate flow (the manipulated variable) to control the drum level.

It took about 160 minutes for the drum level to reach its steady state value after the feed flow perturbation (see curves of H_D and D in Figure B.2).

The simulation results of both open-loop and closed-loop responses were expected. The simulator appears to provide a reliable representation of a distillation process.

Reference

Ogunnaike, B. A.; and W. H. Ray, "Process dynamics, modeling, and control", Oxford University Press, 1994.

APPENDIX C

Empirical Model Identification

C.1 Model Identification

In dynamic data reconciliation, dynamic models are required to provide redundancy in reconciling raw measurements. In order to reconcile the four measurements (reflux drum level, top temperature, bottom level and bottom temperature) for the control of the distillation column, dynamic models for the four controlled variables were developed. Because of the complexity of the distillation process, it is impractical to derive fundamental models for the four controlled variables. Instead, empirical dynamic models were identified, and used in this work.

For process identification, the simple method using process reaction curves (Marlin, 2000) was employed. Using this approach, the identification of the dynamic models for the distillation column followed four steps. Firstly, the column was allowed to reach steady state. Then, with all control loops opened, a step change for each input variable (manipulated and external disturbance) was introduced, one at the time and in turn, to produce dynamic responses of the column. Then, the output (controlled variable) responses were sampled. The magnitudes of the step changes were 20% of the steady-state values. However, a 10% step change in reboiler heat duty was used in identification of the top temperature model to avoid dry-up of the reboiler. At last, the model parameters were calculated. The process reaction curves to a step change in each input variable are presented in Figure C.1. It should be noted that, in the simulation, the

measured data of the four controlled variables were corrupted by white Gaussian noise having the standard deviation listed in Table C.1.

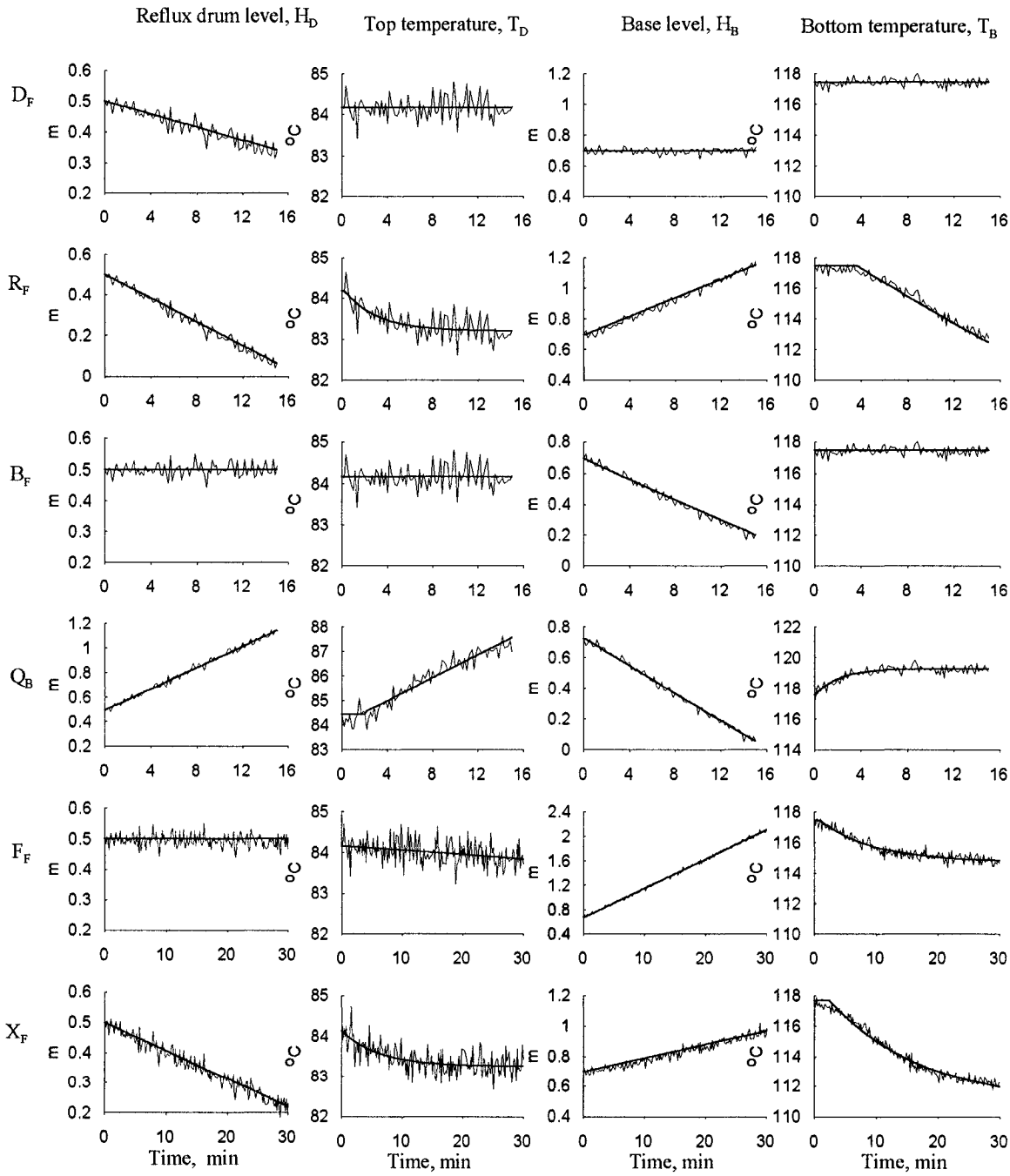

Figure C.1 Process reaction curves to a step change in each input variable.
D_F: Distillate flow R_F: Reflux flow B_F: Bottom flow
Q_B: Reboiler duty F_F: Feed flow X_F: Feed composition
Measured ——— Model predicted ———

Table C.1 Nominal steady-state values and noise levels of measured variables

Measurements	Unit	Steady-state value	Standard deviation
Top temperature (T_D)	°C	84.20	0.25
Bottom temperature (T_B)	°C	117.4	0.25
Reflux drum level (H_D)	m	0.500	0.02
Column base level (H_B)	m	0.700	0.02

Since lower order models are sufficient to capture the essence of process dynamics, pure-integrator-plus-dead-time and first-order-plus-dead-time models were used to describe the input-output behavior of the responses. The mathematical model for pure-integrator-plus-dead-time is written in the Laplace domain.

$$G(s) = \frac{Ke^{-\theta s}}{s} \tag{C.1}$$

where K is the process gain, and θ is the process dead time. On the other hand, the first-order-plus-dead-time model is given by

$$G(s) = \frac{Ke^{-\theta s}}{\tau s + 1} \tag{C.2}$$

where τ is the process time constant. The least-squares technique was employed to determine the model parameters, K, θ and τ in (C.1) and (C.2) (Marlin, 2000). The dynamic models for each reaction curve in Figure C.1, along with the associated model parameters, are presented in Table C.2. It is worth noting that all the variables in Table C.2 are defined in deviation variables.

Table C.2 Empirical dynamic models in the Laplace domain

(Note: θ and τ are in second)

Output / Input	Reflux drum level, m $H_D'(s)$	Top temperature, °C $T_D'(s)$	Base level, m $H_B'(s)$	Bottom temperature, °C $T_B'(s)$
Distillate flow, kmol/h, $D_F'(s)$	$\frac{H_D'(s)}{D_F'(s)} = \frac{-0.1087}{s}$	$\frac{T_D'(s)}{D_F'(s)} = 0$	$\frac{H_B'(s)}{D_F'(s)} = 0$	$\frac{T_B'(s)}{D_F'(s)} = 0$
Reflux flow, kmol/h, $R_F'(s)$	$\frac{H_D'(s)}{R_F'(s)} = \frac{-0.1087}{s}$	$\frac{T_D'(s)}{R_F'(s)} = \frac{-0.0621e^{-14s}}{176s+1}$	$\frac{H_B'(s)}{R_F'(s)} = \frac{0.1228e^{-30s}}{s}$	$\frac{T_B'(s)}{R_F'(s)} = \frac{-1.8344e^{-218s}}{3672s+1}$
Bottom flow, kmol/h, $B_F'(s)$	$\frac{H_D'(s)}{B_F'(s)} = 0$	$\frac{T_D'(s)}{B_F'(s)} = 0$	$\frac{H_B'(s)}{B_F'(s)} = \frac{-0.1382}{s}$	$\frac{T_B'(s)}{B_F'(s)} = 0$
Reboiler duty, MJ/h, $Q_B'(s)$	$\frac{H_D'(s)}{Q_B'(s)} = \frac{0.0038e^{-10s}}{s}$	$\frac{T_D'(s)}{Q_B'(s)} = \frac{0.0297e^{-120s}}{s}$	$\frac{H_B'(s)}{Q_B'(s)} = \frac{-0.0037e^{-10s}}{s}$	$\frac{T_B'(s)}{Q_B'(s)} = \frac{0.0031e^{-28s}}{126s+1}$
Feed flow, kmol/h, $F_F'(s)$	$\frac{H_D'(s)}{F_F'(s)} = 0$	$\frac{T_D'(s)}{F_F'(s)} = \frac{-0.0426e^{-20s}}{s}$	$\frac{H_B'(s)}{F_F'(s)} = \frac{0.1429e^{-10s}}{s}$	$\frac{T_B'(s)}{F_F'(s)} = \frac{-0.1335e^{-60s}}{513s+1}$
Feed composition, mole frac., $X_F'(s)$	$\frac{H_D'(s)}{X_F'(s)} = \frac{-9.0e^{-20s}}{s}$	$\frac{T_D'(s)}{X_F'(s)} = \frac{-18.88e^{-95s}}{372s+1}$	$\frac{H_B'(s)}{X_F'(s)} = \frac{9.1e^{-40s}}{s}$	$\frac{T_B'(s)}{X_F'(s)} = \frac{-114.7e^{-150s}}{917s+1}$

For the multiple-input multiple-output system of the distillation column, the superposition property of the step responses for each output variable was assumed. As a

result, the overall effects of the inputs on the outputs are additive, and the overall dynamic models can be written in the matrix form

$$\begin{bmatrix} H_D'(s) \\ T_D'(s) \\ H_B'(s) \\ T_B'(s) \end{bmatrix} = \mathbf{G}_U(s) \begin{bmatrix} D_F'(s) \\ R_F'(s) \\ B_F'(s) \\ Q_B'(s) \end{bmatrix} + \mathbf{G}_D(s) \begin{bmatrix} F_F'(s) \\ X_F'(s) \end{bmatrix} \tag{C.3}$$

where

$$G_U(s) = \begin{bmatrix} \frac{-0.1087}{s} & \frac{-0.1087}{s} & 0 & \frac{0.0038e^{-10s}}{s} \\ 0 & \frac{-0.0621e^{-14s}}{176s+1} & 0 & \frac{0.0297e^{-120s}}{s} \\ 0 & \frac{0.1228e^{-30s}}{s} & \frac{-0.1382}{s} & \frac{-0.0037e^{-10s}}{s} \\ 0 & \frac{-1.8344e^{-218s}}{3672s+1} & 0 & \frac{0.0031e^{-28s}}{126s+1} \end{bmatrix} \tag{C.4}$$

is the matrix of transfer functions related to the manipulated variables, and

$$G_D(s) = \begin{bmatrix} 0 & \frac{-9.0e^{-20s}}{s} \\ \frac{-0.0426e^{-20s}}{s} & \frac{-18.88e^{-95s}}{372s+1} \\ \frac{0.1429e^{-10s}}{s} & \frac{9.1e^{-40s}}{s} \\ \frac{-0.1335e^{-60s}}{513s+1} & \frac{-114.7e^{-150s}}{917s+1} \end{bmatrix} \tag{C.5}$$

is the matrix of transfer functions related to external disturbances.

C.2 Model Transformation

In the dynamic data reconciliation, since the measurements are sampled at equal-spaced time intervals, it is convenient to use the models in the discrete form rather than in the continuous form as presented by equations (C.3~C.5). The transformation of the continuous models were performed by z-transformation. The transfer function of an output (controlled variable), Y, related to an input, X, (manipulated or external disturbance) in the z-domain is given by (Ogunnaike and Ray, 1994)

$$\frac{Y(z)}{X(z)} = Z[H(s)G(s)] \tag{C.6}$$

where Z represents the z-transform, $G(s) = \frac{Y(s)}{X(s)}$ represents the transfer function in the Laplace domain as listed in Table C.2. H(s) is the zero-order data hold, and it given by

$$H(s) = \frac{1 - e^{-\Delta t s}}{s} \tag{C.7}$$

where Δt is the sampling time interval, and it was 30 seconds for the distillation column. Putting equations (C.1) and (C.7) into (C.6) and taking the z-transformation yield the pure-integrator-with-dead-time model in the discrete form

$$Y_t = \frac{(b_0 - b_1 z^{-1}) z^{-p-1}}{1 - z^{-1}} X_t \tag{C.8}$$

where Y_t is the predicted value of the output variable at time t, X_t is the value of the input variable at time t, z^{-1} is the backward operator such that $X_{t-1} = z^{-1} X_t$, p is the

integer part of the division, $p = \mathrm{INT}(\theta/\Delta t)$, b_0, and b_1 are the coefficients given by $b_0 = K(1-\lambda)\Delta t$, $b_1 = K\lambda\Delta t$ where λ is the remainder of the division $\lambda = REM(\theta/\Delta t)$. Putting equations (C.2) and (C.7) into (C.6) and taking z-transformation give the first-order-with-dead-time model in the discrete form

$$Y_t = \frac{(b_0 - b_1 z^{-1}) z^{-p-1}}{1 - a_1 z^{-1}} X_t \tag{C.9}$$

where p is the integer part of the division, $p = INT(\theta/\Delta t)$, b_0, b_1, and a_1 are the coefficients given by $b_0 = K(1 - e^{-\frac{(1-\lambda)\Delta t}{\tau}})$, $b_1 = K(e^{-\frac{\Delta t}{\tau}} - e^{-\frac{(1-\lambda)\Delta t}{\tau}})$, $a_1 = e^{-\frac{\Delta t}{\tau}}$ where λ is the remainder of the division $\lambda = REM(\theta/\Delta t)$.

Equations (C.8) and (C.9) were used to perform the discretization of the matrices of transfer functions of (C.4) and (C.5). The discrete models are given by

$$\mathbf{HG}_U(z) = \begin{bmatrix} \frac{-9.058\times10^{-4} z^{-1}}{1-z^{-1}} & \frac{-9.058\times10^{-4} z^{-1}}{1-z^{-1}} & 0 & \frac{2.112\times10^{-5}(z^{-1}+0.5z^{-2})}{1-z^{-1}} \\ 0 & \frac{-5.64\times10^{-3}(z^{-1}+0.769z^{-2})}{1-0.8433z^{-1}} & 0 & \frac{2.5\times10^{-4} z^{-5}}{1-z^{-1}} \\ 0 & \frac{1.023\times10^{-3} z^{-2}}{1-z^{-1}} & \frac{-1.152\times10^{-3} z^{-1}}{1-z^{-1}} & \frac{-2.056\times10^{-5}(z^{-1}+0.5z^{-2})}{1-z^{-1}} \\ 0 & \frac{-0.011(z^{-8}+0.35z^{-9})}{1-0.992z^{-1}} & 0 & \frac{4.867\times10^{-5}(z^{-1}+12.5z^{-3})}{1-0.788z^{-1}} \end{bmatrix} \tag{C.10}$$

$$\mathbf{HG}_D(z)=\begin{bmatrix} 0 & \dfrac{-2.5\times10^{-2}(z^{-1}+2.0z^{-2}}{1-z^{-1}} \\ \dfrac{-1.184\times10^{-4}(z^{-1}+2.0z^{-2})}{1-z^{-1}} & \dfrac{-1.227(z^{-4}+0.19z^{-5})}{1-0.9225z^{-1}} \\ \dfrac{7.942\times10^{-4}(z^{-1}+0.5z^{-2})}{1-z^{-1}} & \dfrac{5.06\times10^{-2}(z^{-2}-0.5z^{-3})}{1-z^{-1}} \\ \dfrac{-7.583\times10^{-3}z^{-3}}{1-0.943z^{-1}} & \dfrac{-3.69z^{-6}}{1-0.968z^{-1}} \end{bmatrix} \quad \text{(C.11)}$$

In the z-domain, equation (C.3) can be written as

$$\begin{bmatrix} H'_{D,t} \\ T'_{D,t} \\ H'_{B,t} \\ T'_{B,t} \end{bmatrix} = \mathbf{HG}_U(z)\begin{bmatrix} D'_{F,t} \\ R'_{F,t} \\ B'_{F,t} \\ Q'_{B,t} \end{bmatrix} + \mathbf{HG}_D(z)\begin{bmatrix} F'_{F,t} \\ X'_{F,t} \end{bmatrix} \quad \text{(C.12)}$$

Putting (C.10) and (C.11) into (C.12) yields the model predicted values for each controlled variable for the distillation column at time t

$$H'_{D,t}=\frac{-9.058\times10^{-4}z^{-1}}{1-z^{-1}}D'_{F,t}+\frac{-9.058\times10^{-4}z^{-1}}{1-z^{-1}}R'_{F,t}+\frac{2.112\times10^{-5}(z^{-1}+0.5z^{-2})}{1-z^{-1}}Q'_{B,t}+\frac{-2.5\times10^{-2}(z^{-1}+2.0z^{-2})}{1-z^{-1}}X'_{F,t} \quad \text{(C.13)}$$

$$T'_{D,t}=\frac{-5.64\times10^{-3}(z^{-1}+0.768z^{-2})}{1-0.8433z^{-1}}R'_{F,t}+\frac{2.5\times10^{-4}z^{-5}}{1-z^{-1}}Q'_{B,t}+\frac{-1.184\times10^{-4}(z^{-1}+2.0z^{-2})}{1-z^{-1}}F'_{F,t}+\frac{-1.227(z^{-4}+0.19z^{-5})}{1-0.9225z^{-1}}X'_{F,t} \quad \text{(C.14)}$$

$$H'_{B,t}=\frac{1.023\times10^{-3}z^{-2}}{1-z^{-1}}R'_{F,t}+\frac{-1.152\times10^{-3}z^{-1}}{1-z^{-1}}B'_{F,t}+\frac{-2.056\times10^{-5}(z^{-1}+0.5z^{-2})}{1-z^{-1}}Q'_{B,t}+\frac{7.942\times10^{-4}(z^{-1}+0.5z^{-2})}{1-z^{-1}}F'_{F,t}+\frac{5.06\times10^{-2}(z^{-2}-0.5z^{-3})}{1-z^{-1}}X'_{F,t} \quad \text{(C.15)}$$

$$T'_{B,t}=\frac{-0.011(z^{-8}+0.35z^{-9})}{1-0.992z^{-1}}R'_{F,t}+\frac{4.867\times10^{-5}(z^{-1}+12.5z^{-2})}{1-0.788z^{-1}}Q'_{B,t}+\frac{-7.583\times10^{-3}z^{-3}}{1-0.943z^{-1}}F'_{F,t}+\frac{-3.69z^{-6}}{1-0.968z^{-1}}X'_{F,t} \quad \text{(C.16)}$$

In the application of equations (C.13~C.16), it is more convenient to write them as auto regressive moving average (ARMA) models (Box et al., 1976). It should be noted that

equations (C.14) and (C.16) are higher order if the equations are directly arranged into an ARMA form by multiplying the products of the denominators for each term. In practice, since the higher order models are more sensitive, equations (C.14) and (C.16) were simplified into first-order equations by taking the average values of the denominators for each term in the right-hand side of the equations.

At last, the final equations in the ARMA form can be expressed as

$$H'_{D,t} - H'_{D,t-1} = -9.058\times10^{-4}D'_{F,t-1} - 9.058\times10^{-4}R'_{F,t-1} + 2.112\times10^{-5}Q'_{B,t-1} + 1.06\times10^{-5}Q'_{B,t-2} - 2.5\times10^{-2}X'_{F,t-1} - 5.0\times10^{-2}X'_{F,t-2} \quad (C.17)$$

$$T'_{D,t} - 0.9414T'_{D,t-1} = -5.64\times10^{-3}R'_{F,t-1} - 4.332\times10^{-3}R'_{F,t-2} + 2.5\times10^{-4}Q'_{B,t-5} - 1.184\times10^{-4}F'_{F,t-1} - 2.368\times10^{-4}F'_{F,t-2} - 1.227X'_{F,t-4} - 0.233X'_{F,t-5} \quad (C.18)$$

$$H'_{B,t} - H'_{B,t-1} = 1.023\times10^{-3}R'_{F,t-2} - 1.152\times10^{-3}B'_{F,t-1} - 2.056\times10^{-5}Q'_{B,t-1} - 1.028\times10^{-5}Q'_{B,t-2} + 7.942\times10^{-4}F'_{F,t-1} + 3.971\times10^{-4}F'_{F,t-2} + 5.06\times10^{-2}X'_{F,t-2} - 2.53\times10^{-3}X'_{F,t-3} \quad (C.19)$$

$$T'_{B,t} - 0.9228T'_{B,t-1} = -0.011R'_{F,t-8} - 0.00385R'_{F,t-9} + 4.867\times10^{-5}Q'_{B,t-1} + 6.084\times10^{-4}Q'_{B,t-2} - 7.583\times10^{-3}F'_{F,t-3} - 3.69X'_{F,t-6} \quad (C.20)$$

References

Box, G. P.; and G. M. Jenkins, "Time series analysis, forecasting and control", Holdden-Day, 1976.

Marlin, T. E., "Process control, design processes and control systems for dynamic performance", 2nd Ed., McGraw-Hill, 2000.

Ogunnaike, B. A.; and W. H. Ray, "Process dynamics, modeling, and control", Oxford University Press, 1994.

APPENDIX D

Basic Physical Properties of Benzene and Toluene

Basic properties	Benzene	Toluene
Molecular weight[1]	78.11	92.13
Critical temperature, K[1]	562.05	591.75
Critical pressure, mPa[1]	4.895	4.108
Critical volume, cm^3/mole[1]	256.0	316.0
Liquid density (25 °C), kg/m^3 [1]	876.5	866.9
Surface tension (25 °C), mN/m[1]	28.22	27.93
Vapor enthalpy (25 °C), kJ/mole[1]	82.9	50.5
Liquid enthalpy (25 °C), kJ/mole[1]	49.1	12.4
Heat of vaporization (25 °C), kJ/mole[1]	33.83	38.01
Antoine constant, A[2]	6.01905	6.08436
Antoine constant, B[2]	1204.637	1347.62
Antoine constant, C[2]	220.069	219.787
Vapor heat capacity coefficient, a[3]	74.06×10^{-3}	94.18×10^{-3}
Vapor heat capacity coefficient, b[3]	32.95×10^{-5}	38.00×10^{-5}
Vapor heat capacity coefficient, c[3]	25.20×10^{-8}	27.86×10^{-8}
Vapor heat capacity coefficient, d[3]	77.57×10^{-12}	80.33×10^{-12}
Liquid heat capacity coefficient, a[3]	126.5×10^{-3}	148.8×10^{-3}
Liquid heat capacity coefficient, b[3]	23.4×10^{-5}	32.4×10^{-5}
Vapor viscosity coefficient, a[4]	-0.1509	1.7869
Vapor viscosity coefficient, b[4]	0.25706	0.23566
Vapor viscosity coefficient, c[4]	-8.9797×10^{-6}	-9.3508×10^{-6}
Liquid viscosity coefficient, a[4]	-7.4005	-5.1649
Liquid viscosity coefficient, b[4]	1.1815×10^{3}	8.1068×10^{2}
Liquid viscosity coefficient, c[4]	1.4888×10^{-2}	1.0454×10^{-2}
Liquid viscosity coefficient, d[4]	-1.3713×10^{-5}	-1.0488×10^{-5}

[1] Lide, D. R., "CRC handbook of chemistry and physics", 81st edition, CRC press, 2000-2001.

[2] Boublik, T.; V. Fried; and E. Hala, "The vapor pressure of pure substances", Elsevier, 1984.

[3] Felder, R.; and R. W. Rousseau, "Elementary principles of chemical processes", Third edition, John Wiley & Sons, 2000.

[4] Yaws, C.L., "Handbook of transport properties data, Gulf publishing, 1995.