การปรับให้สอดคล้องพร้อมกับการค้นหาความผิดพลาดอย่างเห็นได้ชัดของข้อมูล โดยใช้วิธีโรบัสฟังก์ชันสำหรับถังปฏิกรณ์ไนล่อน 6

นายจักรพล วีระวงษ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมเคมี ภาควิชาวิศวกรรมเคมี
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2550

SIMULTANEOUS DATA RECONCILIATION AND GROSS ERROR DETECTION USING A ROBUST FUNCTION METHOD FOR NYLON 6 REACTOR

Mr. Jakapon Veeravong

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Chemical Engineering
Department of Chemical Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2007

Thesis Title	SIMULTANEOUS DATA RECONCILIATION AND GROSS ERROR DETECTION USING A ROBUST FUNCTION METHOD FOR NYLON 6 REACTOR
By	Mr. Jakapon Veeravong
Field of Study	Chemical Engineering
Thesis Principal Advisor	Soorathep Kheawhom, Ph.D.

Accepted by the Faculty of Engineering, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

.....Dean of the Faculty of Engineering
(Associate Professor Boonsom Lerdhirunwong, Dr.Ing.)

THESIS COMMITTEE

..... Chairperson
(Assistant Professor Montree Wongsri, D.Sc.)

..... Thesis Principal Advisor
(Soorathep Kheawhom, Ph.D.)

..... External Member
(Assistant Professor Worapon Kiatkittipong, D.Eng.)

..... Member
(Assistant Professor Amornchai Arpornwichanop, D.Eng.)

จักรพล วีระวงษ์ : การปรับให้สอดคล้องพร้อมกับการค้นหาความผิดพลาดอย่างเห็นได้ชัดของข้อมูลโดยใช้วิธีโรบัสฟังก์ชันสำหรับถังปฏิกรณ์ไนล่อน 6. (SIMULTANEOUS DATA RECONCILIATION AND GROSS ERROR DETECTION USING A ROBUST FUNCTION METHOD FOR NYLON 6 REACTOR) อ. ที่ปรึกษาวิทยานิพนธ์หลัก: ดร. สุรเทพ เขียวหอม, 77 หน้า.

ออนไลน์ออฟติไมเซชันเป็นเครื่องมือที่ทรงพลังในการเพิ่มผลผลิตและลดการใช้ทรัพยากรในอุตสาหกรรมเคมี เนื่องจากมันสามารถช่วยให้กระบวนการมีการดำเนินการอยู่ในสภาวะที่ดีที่สุด ดังนั้นในปัจจุบันออนไลน์ออฟติไมเซชันจึงได้รับความสนใจอย่างยิ่ง โดยในการทำออนไลน์ออฟติไมเซชันนั้นจะต้องแก้ปัญหาออฟติไมเซชัน 3 ปัญหาด้วยกันคือ การปรับให้สอดคล้องของข้อมูล การประมาณค่าพารามิเตอร์ และการออฟติไมซ์ในทางเศรษฐศาสตร์ โดยค่าที่ได้จากการวัดตัวแปรกระบวนการจะถูกใช้เพื่อหาค่าที่แท้จริงของตัวแปรกระบวนการ อย่างไรก็ตามในการวัดตัวแปรกระบวนการจะมีความผิดพลาดอย่างเห็นได้ชัดและความผิดพลาดแบบสุ่ม ซึ่งความผิดพลาดทั้งสองแบบนี้ต้องถูกกำจัดออกไปในขั้นตอนการปรับให้สอดคล้องของข้อมูล ในขณะเดียวกันค่าตัวแปรกระบวนการที่แท้จริงต้องเป็นไปตามกฎการอนุรักษ์มวลและพลังงานด้วย ในงานวิจัยนี้เรานำการปรับให้สอดคล้องของข้อมูลมาประยุกต์ใช้ในถังปฏิกรณ์แบบวีเคคอลัมน์สำหรับกระบวนการผลิตไนล่อน 6 ในอุตสาหกรรมที่สภาวะคงตัว เราศึกษาระเบียบวิธีในการทำการปรับให้สอดคล้องของข้อมูล 3 ระเบียบวิธีด้วยกันคือ Contaminated Normal, Lorentzian distribution function และ Hampel's redescending M-estimator เพื่อเปรียบเทียบประสิทธิภาพ นอกจากนั้นยังได้เปรียบเทียบระเบียบวิธีเหล่านี้ในกรณีที่มีบางกระแสไม่มีการวัดด้วย ซึ่งผลที่ได้รับจากการปรับให้สอดคล้องของข้อมูลของแต่ละระเบียบวิธีแสดงให้เห็นว่าทั้ง 3 ระเบียบวิธีสามารถใช้กับกระบวนการนี้ได้ดีพอๆกัน เพราะสามารถรองรับกับข้อมูลการวัดของกระบวนการที่มีทั้งความผิดพลาดอย่างเห็นได้ชัดและความผิดพลาดแบบสุ่มได้ ระเบียบวิธีเหล่านี้ยังสามารถทำงานได้ดีในกรณีที่มีบางกระแสไม่ได้วัดอีกด้วย

ภาควิชา.....วิศวกรรมเคมี..... ลายมือชื่อนิสิต.....จักรพล วีระวงษ์.....
สาขาวิชา.....วิศวกรรมเคมี..... ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก.....สุรเทพ เขียวหอม.....
ปีการศึกษา.....2550.....

4970251621 : MAJOR CHEMICAL EHGINEERINGR
KEY WORD : DATA RECONCILIATION / NYLON 6 / REAL TIME ONLINE OPTIMIZATION / ROBUST FUNCTION METHOD

JAKAPON VEERAVONG : SIMULTANEOUS DATA RECONCILIATION AND GROSS ERROR DETECTION USING A ROBUST FUNCTION METHOD FOR NYLON 6 REACTOR. THESIS PRINCIPAL ADVISOR: SOORATHEP KHEAWHOM, Ph.D., 77 pp.

On-line optimization is a powerful method for economic improvement and resource reduction in chemical industries, because it allows a process to be operated near its optimum operating condition. Thus, it is currently receiving increasing attention. In order to perform online optimization, sequence solving of three optimization problems which are data reconciliation, parameter estimation, and economic optimization are required. Normally, process measurements are used to determine the actual state of the process. However, these measurements usually contain random as well as gross errors. Both random and gross errors should be eliminated in data reconciliation step. Further, the reconciled data must satisfy process material and energy balances. In this work, we apply data reconciliation in VK column reactors for industrial nylon 6 production process at steady state condition. Three data reconciliation algorithms: Contaminated Normal, Lorentzian distribution function and Hampel's redescending M-estimator, are investigated to compare the performances. We also compare these algorithms under the conditions where some process streams are unmeasured. The result shows that all algorithms perform well in this process. Because it can support with process measurements contain both gross and random errors. Moreover, these algorithms also work well in the case where unmeasurement process streams exist.

Department: ...Chemical Engineering... Student's signature: Jakapon Veeravong
Field of study: .Chemical Engineering... Principal Advisor's signature: Soorathep Kheawhom
Academic year: .....2007.....

ACKNOWLEDGEMENTS

The author would like to sincerely thanks and express his sincere gratitude to his advisor, Dr. Soorathep Kheawhom, for his supervision, encouraging guidance, advice, discussion and helpful suggestions throughout the course of this Master Degree study. Furthermore, he is also grateful to Assistant Professor Montree Wongsri, Assistant Professor Amornchai Arpornwichanop, and Assistant Professor Worapon Kiatkittipong for serving as chairperson and member of thesis committees, respectively.

Many thanks to research fund (TRF) for financial supporting this work, Control and Systems Engineering Research Centre members, friends, and all those who encouraged his over the years of his study.

Most of all, the author would like to express the highest gratitude to his family for their love, inspiration, encouragement and financial support throughout this study.

CONTENTS

PAGE

CHAPTER PAGE

CHAPTER PAGE

LIST OF TABLES

PAGE

PAGE

LIST OF FIGURES

PAGE

PAGE

NOMENCLATURES

A_i^0 = Pre-exponential factor for uncatalyzed forward reaction *i* (kg/mol-s)

A_i^c = Pre-exponential factor for catalyzed forward reaction *i* (kg/mol-s)

a_H, b_H, c_H = Turning constants of Hampel' s redescending M-estimator

b_{CN} = Turning constants of Contaminated Normal distribution function

C_c = Turning constants of Cauchy distribution function

C_F = Turning constants of Fair distribution function

C_L = Turning constants of Lorentzian distribution function

C_{Lo} = Turning constants of Logistic distribution function

C_i = Outlet concentration of component *i* (mol/kg)

$C_{i,in}$ = Inlet concentration of component *i* (mol/kg)

E_i^0 = Activation energy for uncatalyzed forward reaction *i* (J/mol)

E_i^c = Activation energy for catalyzed forward reaction *i* (J/mol)

F_{in} = Total mass flow rate of inlet stream (mol/kg)

$F_{i,in}$ = Inlet mass flow rate of component *i* (mol/kg)

$F_{i,v}$ = Mass flow rate of component *i* in vapor phase (kg/hr)

ΔH_i = enthalpy of reaction *i* (J/mol)

K_i = Equilibrium constant for reaction *i*

k_i = Forward rate constant for reaction *i* (kg/mol-s)

$k_i^{'}$ = Reverse rate constant for reaction *i* (kg/mol-s)

M = Total mass in the reactor (kg)

P_n = Polymer chain with degree of polymerization *n*

R = Perfect-gas constant (J/mol-K)

R_i = Rate equation for reaction *i*

ΔS_i = Entropy of reaction *i* (J/mol-K)

T = Temperature (K)

y^* = Vector of true value

GREEK LETTERS

σ_i = Standard deviation

ε = Vector of random error

δ = Vector of gross error

η = Probability of gross error measurements

SUBSCRIPTS

$B-ACA$ = Nylon-6 repeat segments

CD = Cyclic Dimer

CL = Caprolactam

in = Inlet

P_i = Amino Caproic Acid

$T-COOH$ = Terminal carboxylic acid

$T-NH_2$ = Terminal amine

W = Water

ACRONYM

CN = Contaminated Normal

CSTR = Continuous stirred tank reactor

DCS = Distributed Control System

DE = Different Evolution

DR = Data Reconciliation

EAs = Evolutionary algorithms

GED = Gross Error Detection

RTO = Real-time Optimizatiom

WLS = Weighted Least Square

CHAPTER I

INTRODUCTION

This chapter is an introduction of this research. We first introduce importance and reasons for research, research objectives, scopes of research, contributions of research, research procedures, and research contents.

Chapter organization is as follows:

1.1 – Importance and reasons for research

1.2 – Research Objective

1.3 – Scopes of research

1.4 – Contributions of research

1.5 – Research procedures

1.6 – Research Contents

1.1 Importance and Reasons for Research

In a very dynamic market with globalization, chemical process industrial are increasingly compelled to operate profitably. The increasing competition and stringent product requirements decrease profit margin. Thus, plant operations must be optimized dynamically in order to cope the changing markets conditions and to reduce the operating cost. Hence, the importance of real-time on-line optimization of an entire plant is rapidly increasing.

Real-time optimization (RTO) refers to evaluation and alteration operating conditions of a process continually in order to maximize the economic productivity of the process. Currently, it is more and more used in the chemical industrial in order to operate a process near its optimum condition by providing real-time computed optimal set-points to the distributed control system (DCS). Typical structure and components of RTO are illustrated in Figure 1. Plant measurements collected via the distributed control system are first checked for steady state operation. If the plant is at

steady state, reconciliation and gross error detection are performed on the measured data, and the process model is updated based on reconciled data. Optimization is then carried out using the updated model along with the economic data and product requirements, to find the new set-points for the operating variables. The new set-points are then passed to the distributed control system for implementing on the plant.

The optimization module heavily relies on accuracy of the process model. Moreover, the quality of the measurement data is crucial for the realization of the optimization results. However, in most cases there exist discrepancies between the model and the real plant and the measurements are contaminated with measurement errors.

Most process models have parameters which have to be estimated from measurement data. To improve the accuracy of the model these parameters must be estimated with measurement data taken directly from the plant. Therefore, process

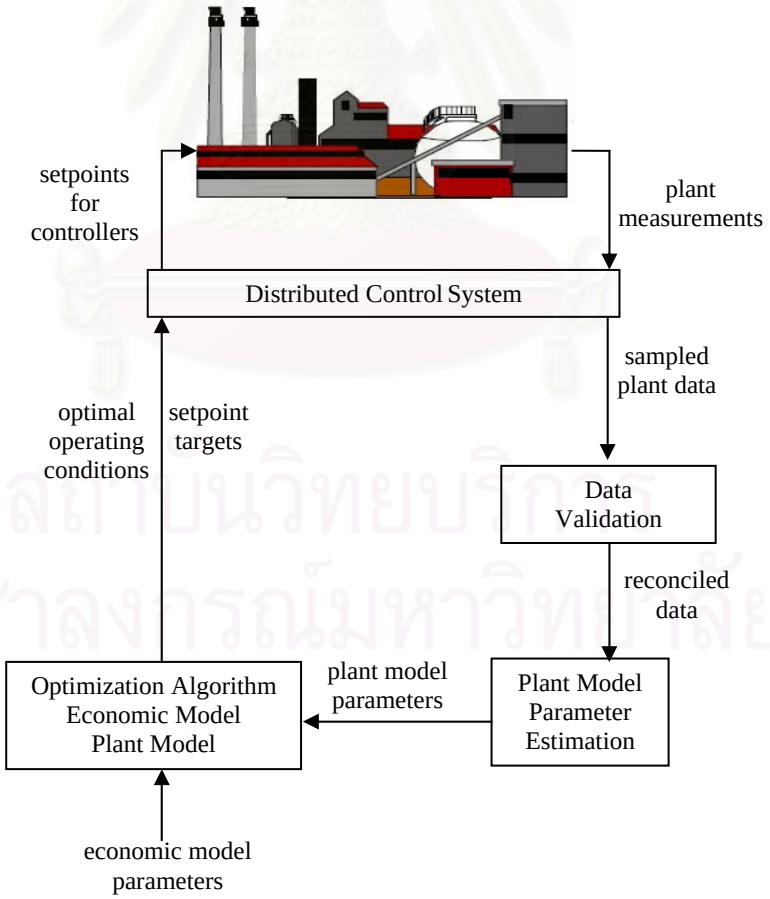


Figure 1.1 Simplified structure of on-line optimization
(Zejun, Ralph, & Thomas, 1995)

measurements are necessary to determine the actual state of the process, and to increase the accuracy of the model. Therefore, measurement data do not only affect the quality of the optimization results but also that of the estimated process model.

Process measurements are inevitably corrupted by errors during the measurement itself, and also during its processing and transmission stages. Generally, measurements often contain random and possibly gross errors as a result of miscalibration or failure of the measuring instruments. These errors should be eliminated before the measurements are used for online optimization.

Data reconciliation is an important step in real time on-line optimization. It adjusts the process measurements with random errors to satisfy the constraints of the system model and provides estimates for unmeasured variables and process parameters, which are used in the consecutive economic optimization step. Therefore data reconciliation has to be performed in order to identify and rectify measurement errors. Reconciled process data is used to specify the current status of the plant model and for estimation of the model parameters for plant-model matching. Most elimination of the less frequent gross errors is achieved by gross error detection. Therefore, data reconciliation and gross error detection are a way to improve the quality of the measurements.

Generally, to improve the quality of measured data includes three steps: steady state identification, gross error detection and data reconciliation. In this research, we focus on the simultaneous data reconciliation and gross error detection strategies to reduce the time required in data validation. We implement simultaneous data reconciliation and gross error detection strategies to industrial nylon 6 production process at steady state condition. We study the performance of each available technique that suitable for industrial nylon 6 production process and compare these algorithms under the conditions where some process streams are unmeasured.

1.2 Research Objective

The objectives of our research are to take advantage of information redundancy on a process to make a cross-check of real time process measurements by

combining data reconciliation with gross error detection, and to apply the developed methodology in a case study of industrial nylon 6 production process.

1.3 Scopes of research

1. Simulation of nylon 6 production process in Continuous stirred tank reactor (CSTR) is studied.

2. An application of data reconciliation with gross error detection in VK column reactors for industrial nylon 6 production process is considered.

3. The data reconciliation problem is formulated as the optimization problem by Weighted Least-Square and Robust function methods. The robust distribution functions that we studied are as follows:

- Contaminated normal distribution.
- Lorentzian distribution.
- Hampel's redescending M-estimator.

4. The data reconciliation problem formulated is solved by in-house optimizer. (Deferential Evolution method)

5. To compare performances each algorithm, take solution to obtain each algorithm compare to true value of each variable at steady state condition. We also compare these algorithms under the conditions where some process streams are unmeasured.

1.4 Contributions of research

The contributions of this research are as follows:

1. An efficient system to estimate the current status of process variables and unmeasured variables of industrial nylon 6 production process.

2. The formalizations of knowledge in applying simultaneous data reconciliation and gross error detection in industrial nylon 6 production process.

1.5 Research procedures

1. Firstly, relevant information regarding nylon 6 production process and data reconciliation is reviewed.

2. A process description of an industrial nylon 6 production process is thoroughly studied.

3. Simulation of nylon 6 production process in Continuous stirred tank reactor (CSTR) at steady state condition.

4. Industrial process operation data are collected.

5. Various different data reconciliation with gross error detection techniques is performed. The data reconciliation with gross error detection problem is formulated as the optimization problem constituted by an objective function that corresponds to maximum likelihood function formed from the probability distribution function of the measured variable. The constraints are mass and energy balances, separation rules, and thermodynamic behaviors.

6. The data reconciliation with gross error detection problem formulated is solved by in-house optimizer.

7. To compare performance each algorithms, take solution to obtain each algorithm compare to true value of each variable at steady state condition.

8. Finally, we conclude our research and write thesis.

1.6 Research Contents

This thesis is divided into five chapters.

Chapter I is an introduction to this research. This chapter consists the importance and reasons for research, objectives of research, scopes of research, contributions of research, and research procedures.

Chapter II is literature reviews related to data reconciliation with gross error detection.

Chapter III cover some background information of data reconciliation, gross error detection, simultaneous data reconciliation and gross error detection, benefit from data reconciliation and gross error detection, and algorithm for solve data reconciliation problem.

Chapter IV Take data reconciliation with gross error detection techniques are performed in industrial nylon 6 production process and the other case studied. And results to obtain data reconciled each algorithms are presented in this Chapter.

Chapter V presents the conclusion of this research and makes the recommendations for future work.

This is follow by:

References

Appendix A: Reaction and Kinetics for Nylon 6 Polymerization

Appendix B: Reconciliation Solution for the Measured Variables in All Cases

CHAPTER II

LITERATURE REVIEW

For more than twenty years, reconciliation problem has received consideration in the literature. Kuehn and Davidson (1961) introduced a data reconciliation method based on linear process models by use Lagrange multipliers to solve for optimal adjustments to measurements for the case when either all or none of component flow rates are measured. The method was improved, for instance, by Nogita [1972] and Mah et al. [1976] introducing new methods for detection of gross errors in measurements. The enhancement of data reconciliation towards nonlinear models was published by Crowe [1986] using matrix projection. In a later article, Crowe gave a survey on nonlinear data reconciliation and the challenges for its further development [1996].

The method of data reconciliation can be corrupted by faulty sensors or improper process models. Therefore, gross errors have to be dealt with additionally. The original method to cope with gross errors considers a sequential approach, in which measurements are eliminated after being detected as afflicted with gross errors and the data reconciliation algorithm is restarted.

In the early 1990s, Tjoa and Biegler (1991) introduced a method which simultaneously reconciles the data and detects the gross errors by combining the treatment of small measurement errors and gross errors into a so-called contaminated Gaussian objective function instead minimize an objective function that is constructed using maximum likelihood principle to construct a new distribution function, which takes into account both contributions from random and gross errors. The advantages of minimizing this objective function are that it gives unbiased estimates in the presence of gross errors and that simultaneously a gross-error detection test can be constructed based on their distribution functions without the assumption on the linearity of the constraints. Furthermore, the structure of this objective function can be exploited under certain conditions. Thus, efficient nonlinear programming strategies,

similar to the hybrid SQP method introduced by Tjoa and Biegler in 1991 for least squares objective functions, can also be developed. The effectiveness of this strategy is demonstrated on nonlinear example problems.

Johnson and Kramer (1995) reported the feasibility and better performance of the robust estimators as the objective function in the data reconciliation problem especially when the data contain gross errors. These robust functions are the Lorentzian distribution. This approach does not divide the sensors into "normal" and " gross error" classes, but uses all of the data in the rectification. In this manner, the conventional assumption of no sensor bias is avoided, and both random errors (noise) and systematic errors (gross errors) are removed simultaneously. This method is demonstrated on data from a simulated flow network and a simulated heat-exchanger network. And briefly discussed the theoretical evaluation of algorithms using the influence function.

Chen, Pike, Hertwig and Hoppe (1998) studied optimal implementation of on-line optimization for Monsanto sulfuric acid contact plant. In data validation step, simultaneous gross error detection and data reconciliation algorithms are used to detect and rectify the gross errors in measurements. These algorithms are measurement test method using a normal distribution, Tjoa-Biegler's method using a contaminated Gaussian distribution, and robust method using robust distribution functions (Lorentzian distribution, Fair distribution). In summary, the evaluation of influence functions for the probability distributions shows that the contaminated Gaussian and Lorentzian distributions have influence functions that are relatively insensitive to gross errors. Methods based on the contaminated Gaussian distribution should have the best performance for reconciling measurements when moderate size gross errors are present (rang $3\sigma - 30\sigma$), and methods using the Lorentzian distribution should be more effective for very large gross errors.

Özyurt and Pike (2004) compare different objective functions with the contaminated Gaussian function regarding their effectiveness in detecting gross errors of simultaneous procedures for data reconciliation and gross error detection is established. These procedures depending on the results from robust statistics reduce the effect of the gross errors. They provide comparable results to those from methods such as modified iterative measurement test method (MIMT) without requiring an

iterative procedure. The comparative results of the introduced methods are given for five literature and more importantly, two industrial cases. Methods based on the Cauchy distribution and Hampel's redescending M-estimator give promising results for data reconciliation and gross error detection with less computation.

CHAPTER III

THEORY

The aim of this research is to take advantage of information redundancy on a process to make a cross-check of real time process measurements by combining data reconciliation with gross error detection, and to apply the developed methodology in a case study of industrial nylon 6 production process. Since major roles of data reconciliation are reconciliation of measured process data to satisfy defined constraints. To efficient system to estimate the current status of process variables and unmeasured variables of industrial nylon 6 production process.

In this chapter, that is to say background information of data reconciliation, gross error detection, simultaneous data reconciliation and gross error detection in each techniques, benefit from data reconciliation and gross error detection and algorithm for solve data reconciliation problem.

3.1 Data Reconciliation

3.1.1 Introduction

Process measurements are inevitably corrupted by errors during the measurement, processing and transmission of the measured signal. Using this information without any filtering technique in process control may affect the achievement of optimal plant performance (e.g. quality, yield or due-date) and even could drive the plant to an unsafe situation (Chouaib, 2004). Therefore, data reconciliation (DR) is technique that has been developed to improve the accuracy of measurements by reducing the effect of random error in the data. The principal difference between data reconciliation and other filtering techniques is that data reconciliation explicitly makes use of process model constraints and obtains estimates of process variables by adjusting process measurements so that the estimates satisfy

the constraints (Narasimhan & Jordache, 2000). Thus, data reconciliation is an imperative procedure in control strategy to improve the accuracy of measurements The data provided via data reconciliation are defined as the optimal solution to a constrained least square and maximum likelihood objective function. The optimal estimates of physical properties such as concentration and temperature are employed in control strategy to reduce level of process data corruption and improve process performance, leading to better quality control.

The reconciled estimates are expected to be more accurate than the measurements and, more importantly, are also consistent with the known relationships between process variables as defined by the constraint. In order for data reconciliation to be effective, there should be no gross error either in the measurement or in the process model constraints. Gross error detection is a companion technique to data reconciliation that has been developed to identify and eliminate gross error. Thus, data reconciliation and gross error detection are applied together to improve accuracy of measured data.

Data reconciliation and gross error detection both achieve error reduction only by exploiting the redundancy property of measurement. Typically, in any process the variables are related to each other through physical constraint such as material or energy conservation laws. Given a set of such system constraint, a minimum number of error-free measurements is required in order to calculate all of the system parameters and variables. If there are more measurements than this minimum, then redundancy exists in the measurements that can be exploited. This type of redundancy is usually called *spatial redundancy* and the system of equation is said to be *overdetermined.*

Data reconciliation cannot be performed without spatial redundancy. With no extra measured information, the system is *just determinated* and no correction to erroneous measurements is possible. Further, if fewer variables than necessary to determine the determine the system are measured, the system is underdetermined and the values of some variables can be estimated only through other means or if additional measurements are provided.

A second type of redundancy that exists in measurements is *temporal*

redundancy. This arises due to the fact that measurements of process variables are made continually in time at a high sampling rate, producing more data than necessary to determine a steady state process. If the process is assumed to be in a steady state, then temporal redundancy can be exploited by simply averaging the measurements, and applying steady state data reconciliation to the averaged values.

If the process state is dynamic, however, the evolution of the process state is described by differential equation corresponding to mass and energy balance, which inherently capture both the temporal and spatial redundancy of measured variables. For such a process, dynamic data reconciliation and gross error detection techniques have been developed to obtain accurate estimates consistent with the differential model equations of the process (Narasimhan & Jordache, 2000).

In general, the total error in a measurement, which is the difference between the measured value and the (definitely unknown) value of a variable, can be conveniently represented as the sum of the contributions from two types of errors: random and gross errors.

1. Random errors.

Random errors which are inherent to the measurement process are usually small in magnitude and are most often described by the use of probability distributions. These errors can be caused by a number of different sources such as power supply fluctuation, network transmission and sign conversion noise, changes in ambient conditions, and so on.

2. Non-random errors or gross errors

Gross errors are caused by nonrandom events such as instrument malfunctioning (due to improper installation of measuring devices), miscalibration, wear or corrosion of sensors and so on. The nonrandom nature of these errors implies that at any given time they have a certain magnitude and sign which may be unknown. Thus, if the measurement is repeated with the same instrument under identical conditions, the contribution of a systematic gross error to the measurement value will be the same.

By following good installation and maintenance procedures, it is possible to ensure that gross error are not present in the measurement at least for some time. Gross error caused by sensor miscalibration may occur suddenly at a particular time and thereafter remain at a constant level or magnitude. Other gross error causes such as the wear or fouling of sensors can occur gradually over a period of time and so the magnitude of the gross error increases slowly over a relatively long time period. Therefore, gross errors occur less frequently but their magnitudes are typically larger than those of random errors. (Narasimhan & Jordache, 2000)

3.1.2 Definition Different Objective Functions for Formulate Data Reconciliation Problem is as the Optimization Problem

3.1.2.1 Weighted Least-Square (WLS)

Steady-State Data Reconciliation (SSDR) was first addressed in the pioneer work of Kuehn and Davidson (1961). The authors adjust process data to satisfy mass balance. Therefore, they formulated the SSDR as a Weighted Least-Square (WLS) optimization problem (show as equation 3.1) subject to mass balances (show as equation 3.2).

$$\min_{\hat{y}}\left(\left(y-\hat{y}\right)^{T} Q^{-1}\left(y-\hat{y}\right)\right)=\min_{\hat{y}}\left(a^{T}Q^{-1}a\right)=\min_{\hat{y}}\left(\sum_{i}\left(\frac{\hat{y}_i-y_i}{\sigma_i}\right)^{2}\right) \qquad (3.1)$$

subject to:

$$A\hat{y}=0 \qquad (3.2)$$

where A is the incidence matrix representing steady state mass balance, it is considered that all process variables involved in the mass balance were measured and the measurement do not contain gross errors. The term a is the adjustment vector that is the difference between the measurement vector (y) and estimated vector ($\hat{y}$) and Q is the variance-covariance matrix. The measurement errors follow a normal distribution with zero-mean and a known variance $Q_{ii}=\sigma_i^2$

The square of standard deviation σ_i^2 is the weight on the measurement adjustment *i*. Variables known with a high certainty (low variability) are given a large weight and variables with high variability implies that the measurement is less accurate and received less weight in Data Reconciliation procedure.

The standard deviation of a measurement error plays an important orle in data reconciliation and various other error reduction techniques. Since the true standard deviation is never know, an estimates of standard deviation can be obtained by using a sample standard deviation, according to the follow equation 3.3

$$s = \frac{1}{N-1}\left[\sum_{i=1}^{N}\left(y_i - \bar{y}\right)\right]^{\frac{1}{2}} \tag{3.3}$$

where s is the estimated value of standard deviation, y_i is the *i*th observation and $\bar{y}$ is the arithmetic average of N observations of the same variable.

The effect of random errors on measurement is modeled as additive contributions. The relation between the measured value, true value and random error in the measurement expressed in equation 3.4.

$$y = y^* + \varepsilon \tag{3.4}$$

where y^* is the vector of true value (noise free), and ε is the vector of random error. The random error (ε) usually oscillates around zero. Its characteristics can be described using statistical properties of random variables. Its mean or *expected value* is zero and its variance is given by:

$$\mathrm{var}\left(\varepsilon_i\right) = E\left[\varepsilon_i^2\right] = \sigma_i^2 \tag{3.5}$$

where σ_i is the standard deviation of the measurement error ε_i.

The problem above described can be solved analytically by using Lagrange multipliers as shown in equation 3.6.

$$\hat{y} = y - QA^T\left(AQA^T\right)^{-1}Ay \tag{3.6}$$

3.1.2.2 Maximum Likelihood Estimation (MLE).

If the measurement error distribution follows a normal distribution, the Data Reconciliation problem can be posed as a Maximum Likelihood Estimation (MLE) problem, where the probability of the estimated (reconciled) process variables (y) is maximized given the measurement set (y) as shown in equation 3.7.

$$\max_{\hat{y}} P\{\hat{y}/y\} \tag{3.7}$$

According to Bayes' theorem, the probability of the process variables given the measurements can be written in terms of the probability of the measurements given the reconciled process variables, the probability density function of the process variables $P\{\hat{y}\}$ and the probability density function of the measurements $P\{y\}$.

$$\max_{\hat{y}} P\{\hat{y}/y\} = \max_{\hat{y}} \frac{p\{y/\hat{y}\}\, p\{\hat{y}\}}{p\{y\}} \tag{3.8}$$

The denominator term (independent of $\hat{y}$) acts as normalizing constant and does not need to be further considered for optimization. The first term in the numerator represents the probability density of the measurements given the reconciled process variables, $\hat{y}$, which is the distribution of the measurements errors P($\hat{y}$ -y). Finally, $P\{\hat{y}\}$ is a binary assumption, that is equal to 1 if the constraints are satisfied (under this assumption the $P\{\hat{y}\}$ term converted to a set of constraints and the original problem is converted to a constrained optimization) and equal to 0 otherwise.

$$P(\hat{y}-y) = P(\varepsilon) \square N(0,Q) \tag{3.9}$$

If sensor errors are independents the product of this probability over all sensors yields to:

$$P\{\hat{y}/y\} = \prod_i \exp\left\{-\frac{1}{2}\left(\frac{\hat{y}_i - y_i}{\sigma_i}\right)^2\right\} = \exp\left\{-\frac{1}{2}\sum_i\left(\frac{\hat{y}_i - y_i}{\sigma_i}\right)^2\right\} \tag{3.10}$$

Taking the negative logarithm of the maximization of the objective function

represented in equation 3.10 results in the minimization of the conventional WLS formulation as is shown in equation 3.1. The symmetric and positive definite matrix Q contains the variance-covariance elements of the measurement errors and thus quantifies the uncertainty in each measured value. Then the success of Data Reconciliation technique relies on the hypothesis that the error is normally distributed and on the evaluation of matrix Q.

3.2 Gross Error Detection (GED)

The technique of data reconciliation crucially depend on the assumption that d values, only random error are present in the data and systematic errors either in the measurement or the model equation are not present. If this assumption is valid, reconciliation can lead to large adjustments being made to the measured values, and the resulting estimates can be very inaccurate and even infeasible. Thus it is important to identify such systematic or gross error before the final reconciled estimated are obtained.

There are two major types of gross errors. One is related to the instrument performance and includes measurement bias, drifting, miscalibration, and total instrument failure. The other is constraint model-related and includes unaccounted loss of material and energy resulting from leaks from process equipment or model inaccuracies due to inaccuracies parameters.

Usually gross errors are associated with sensor faults. In Figure 3.1, illustrates graphically the most common types of instrument; bias, complete failure, drifting, and precision degradation.

Various techniques have been designed for the detection and elimination of these two types of gross errors: statistical tests approaches. Any comprehensive gross error detection strategy should preferably processes the following capabilities:

- Ability to detect the presence of one or more gross error in the data (the detection problem)
- Ability to identify the type and location of the gross error (the identification problem)

- Ability to locate and identify multiple gross error which may be present simultaneous in the data (the multiple gross error identification problem)

- Ability to estimate the magnitude of the gross error (the estimation problem)

A number of statistical tests are derived from this basic statistical principle and are able to detect gross errors. But not all statistical test are able to identify different types and location of gross errors. Some basic statistical test are able to detect only measurement error (biases). Other statistic test can only detect process model error or leaks. On the other hand, the generalized likelihood ratio test, which is derived from maximum likelihood estimation principle in statistics, can be used to detect both instrument problems and process leaks.

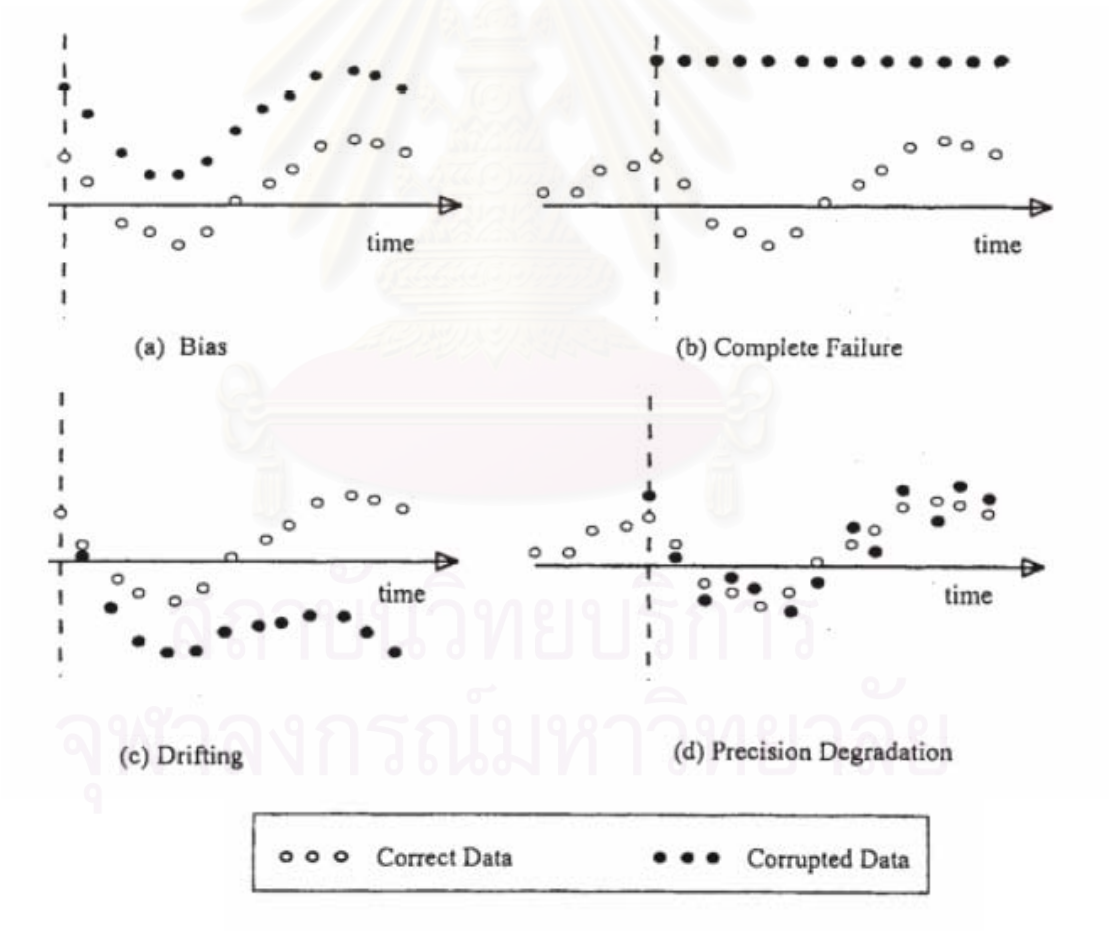


Figure 3.1 Types of gross errors (Narasimhan & Jordache, 2000)

3.3 Simultaneous Data Reconciliation and Gross Error Detection

The process data from a distributed control system is subject to random and gross error, and the gross error must be detected and rectified before the data is used to estimate plant parameters. Simultaneous gross error detection and data reconciliation algorithms are used to detect and rectify the gross errors in measurements. Two main approaches can be adopted for such purpose: the Bayesian approach and Robust approach.

3.3.1. Bayesian approach

Tjoa and Biegler (1991) have proposed a contaminated Gaussian distribution function to describe the measurement errors. A measurement is subject to either random or gross error. The two possible outcomes are: G = {Gross error occurred} with prior probability η and R = {Random error occurred} with prior probability $(1-\eta)$. Therefore, the distribution of a measurement error is:

$$P(y_i / x_i) = (1-\eta) P(y_i / x_i, R) + \eta P(y_i / x_i, G) \qquad (3.11)$$

where $P(y_i / x_i, R)$ is the probability distribution of a random error and

$P(y_i / x_i, G)$ is the probability distribution of a gross error.

η is the probability of gross error measurements

It was assumed that the random errors are normally distributed with a zero mean and a known variance σ_i^2. Also, it was assumed that the gross errors are subject to a contaminated normal distribution which has a zero mean and larger variance $(b\sigma_i)^2, (b \geq 1)$. If the measurement errors are independent of each other, then the likelihood function (or joint probability function) for all measurement, are the products of the distributions for individual measurement, and the measurement errors are estimated by minimizing the negative logarithm of the joint probability density

function. This gives the objective function used with the constraints of equation 3.2 for Tjoa-Biegler's method as (Ozyurt and Pike,2004):

$$\max \prod_i P_i = \max \prod_i \left\{ (1-\eta) \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(y_i - x_i)^2}{2\sigma_i^2} \right) + \eta \frac{1}{\sqrt{2\pi} b \sigma_i} \exp\left(-\frac{(y_i - x_i)^2}{2b^2\sigma_i^2} \right) \right\} \tag{3.12}$$

or

$$\min \; -\sum \ln \left\{ (1-\eta) \exp\left(-\frac{(y_i - x_i)^2}{2\sigma_i^2} \right) + \frac{\eta}{b} \exp\left(-\frac{(y_i - x_i)^2}{2b^2\sigma_i^2} \right) \right\} + \sum_i \ln\left(\sqrt{2\pi}\sigma_i \right) \tag{3.13}$$

where: b is the ratio of the standard deviation of gross errors to that of random errors.

However, the gross error distribution is usually a posteriori information which, may lead to a biased estimation. If the gross error term is higher than the random term for a particular measurement a gross error is identified. In consequence, this approach can only be used if the gross error distribution is known a priori.

3.3.2. Robust approach

The basic idea of robust estimation is to build a robust distribution function ρ that is asymptotic to the normal distribution or any pre-assumed rigorous distribution function that describes the distribution pattern of measurement errors under some ideal assumptions. The estimator (mean or variance) determined by the robust distribution is insensitive to extreme observations and yet maintains a high efficiency (lower dispersion) (Chen, Pike, Hertwig and Hoppe, 1998).

(Huber (1981); Romagnoli and Sanchez (2000)) attempts to make the estimation insensitive in front of the presence of gross errors. The weighted squared residual of the DR formulation is replaced by another function of the residual as shown in equation 3.14. (Chouaib ,2004)

$$\min \sum_i \rho\left(\frac{y_i - x_i}{\sigma_i} \right) \tag{3.14}$$

where $\rho\left(\frac{y_i - x_i}{\sigma_i}\right)$ is usually selected as a convex function in order to ensure that the solutionis unique, the influence function is the derivative of ρ with respect to the process variable measurements. This Robust Estimator will give an unbiased estimate when the gross error follows a previously known distribution and still behaves well if they are deviations from ideal situation. Thus, this influence function compensates for the effects that have the residuals on the estimations, given a weight of zero to high value residuals. The crucial step in the Robust Estimation is the choice of these influence functions: different pre-selected choices of the influence function deals to estimations with different robustness.

The ρ function have been studied previous as follows: (Ozyurt and Pike, 2004)

1. Normal distribution.

$$\frac{1}{2}\varepsilon_i^2 \tag{3.15}$$

2. Contaminated normal distribution function.

$$-\ln\left\{(1-\eta)\exp\left(-\frac{\varepsilon_i^2}{2}\right)+\frac{\eta}{b_{CN}}\exp\left(-\frac{\varepsilon_i^2}{2b_{CN}^2}\right)\right\} \tag{3.16}$$

3. Cauchy distribution function.

$$C_c^2 \ln\left(1+\frac{\varepsilon_i^2}{C_c^2}\right) \tag{3.17}$$

4. Logistic distribution function.

$$2\ln\left(1+\exp\left(\frac{\varepsilon_i}{c_{Lo}}\right)\right)-\left(\frac{\varepsilon_i}{c_{Lo}}\right) \tag{3.18}$$

5. Lorentzian distribution function.

$$-\frac{1}{1+\left(\varepsilon_i^2/2c_L^2\right)} \tag{3.19}$$

6. Fair distribution function.

$$2c_F^2\left[\frac{|\varepsilon_i|}{c_F}-\ln\left(1+\frac{|\varepsilon_i|}{c_F}\right)\right] \tag{3.20}$$

7. Hampel' s redescending M-estimator

$$\begin{aligned}
&\frac{1}{2}\varepsilon_i^2, \qquad 0\le\frac{|\varepsilon_i|}{c_F}\le a_H\\
&a_H|\varepsilon_i|-\frac{1}{2}a_H^2, \quad a_H<|\varepsilon_i|\le b_H\\
&a_Hb_H-\frac{a_H^2}{2}+\left(c_H-b_H\right)\frac{a^2}{2}\left[1-\left(\frac{c_H-|\varepsilon_i|}{c_H-b_H}\right)^2\right], \quad b_H<|\varepsilon_i|\le c_H\\
&a_Hb_H-\frac{1}{2}a_H^2+\left(c_H-b_H\right)\frac{1}{2}a^2, \quad c_H<|\varepsilon_i|
\end{aligned} \tag{3.21}$$

where $\varepsilon_i=\left(y_i-x_i\right)/\sigma_i$ is standard error

$C_c, C_{Lo}, C_L, C_F, C_H$ are turning constants of each distribution function.

3.4 Benefit from Data Reconciliation and Gross Error Detection

Development of a data reconciliation and gross error detection package for a system and its practical implementation is a difficult and costly task and can not be justified without its benefits for a particular industrial application. The justification for data reconciliation and gross error detection may come from the many important applications for improving process performance shown in Figure 3.2 which requires accurate data for achieving expected benefits as outlined below:

1. A direct application of data reconciliation is in evaluating process yields or in assessing consumption of vitalities in different process units. Reconciled values provide more accurate estimates as compared to the use of law measurements.

2. Applications such as simulation and optimization of existing process equipment rely on a model of the equipment. The models usually contain parameters which have to be estimated from plant data. This is also known as model tuning, for which accurate data is essential. The use of erroneous measurements in model tuning can give rise to incorrect model parameters which can nullify the benefits achievable though optimization.

3. Data reconciliation can be very useful in scheduling maintenance of process equipment. Reconciled data can be used to accurately estimate key performance parameters of process equipment.

4. Many advanced control strategies such as model-based control or inferential control require accurate estimates of controlled variables. Dynamic data reconciliation techniques can be used to derive accurate estimates for better process control.

5. Gross error detection not only improves the estimation accuracy of data reconciliation procedures but is also useful in identifying instrumentation problems which require special maintenance and correction. Incipient detection of gross error can reduce maintenance costs and provide a smoother plant operation. These methods can also be extended to detect faulty equipment.

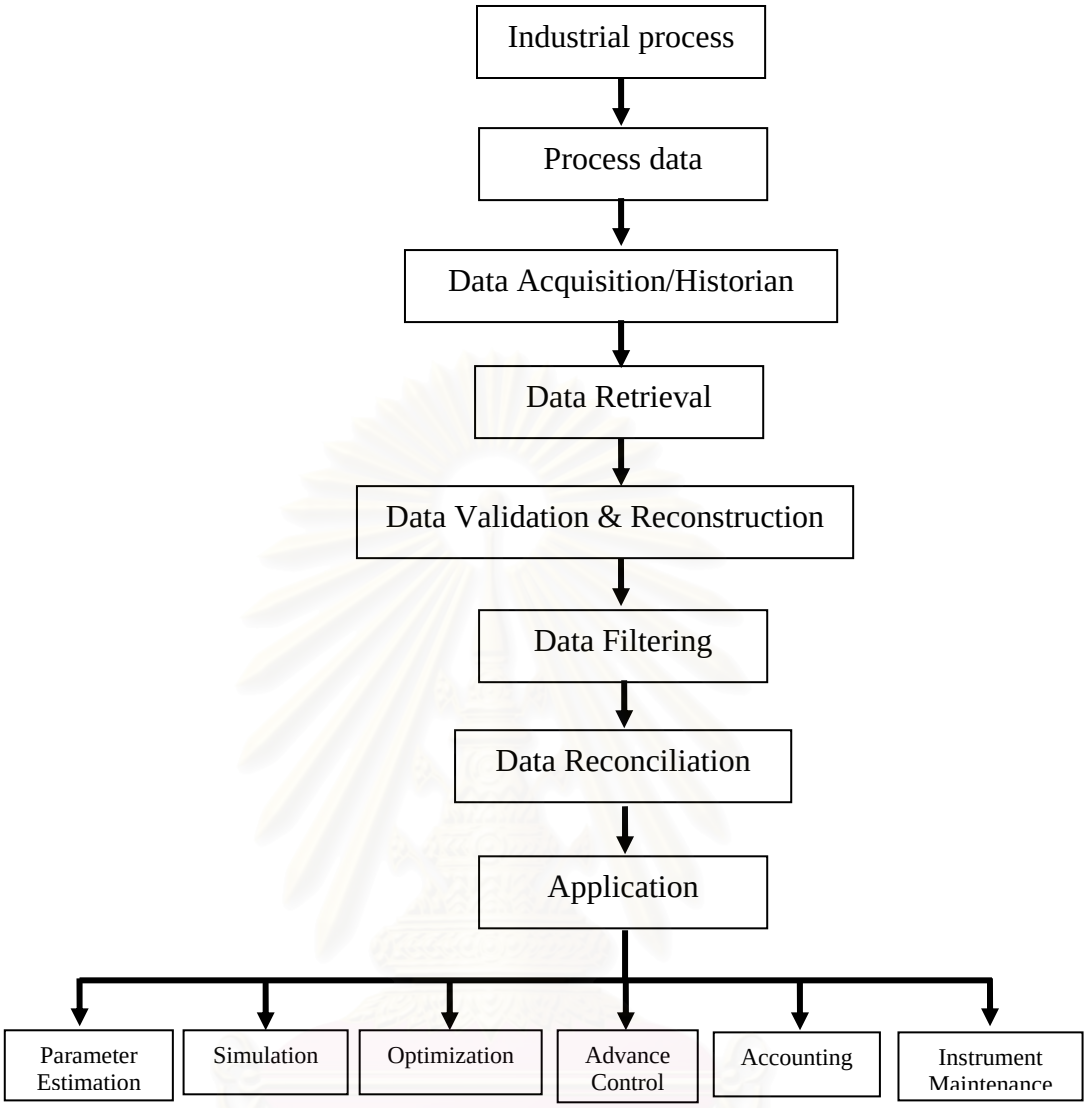


Figure 3.2 Online data collection and conditioning system (Narasimhan & Jordache, 2000)

3.5 Algorithm for Solve Data Reconciliation Problem

In the optimization process of a different task the method of first choice will usually be a problem specific heuristic. Different Evolution (DE) algorithm is a stochastic optimization method minimizing an objective function that can model the problem's objective while incorporation constraints. The algorithm mainly has three advantages: finding the true global minimum regardless of the initial parameter values, fast convergence, and using a few control parameters.

3.5.1 Differential Evolutionary Algorithm

DE was first introduced by Storn & Price. As it is typical for evolutionary algorithms (EAs), DE does not require any prior knowledge of the search space, nor of the derivative information. It is a very simple population based, stochastic optimization algorithm which is very powerful and robust at the same time. Figure 3.3 shows the flowchart of DE. The algorithm starts by generating a randomly distributed initial population of N vectors. Mutation and recombination is then performed on each vector X_i of the generated population in order to create a trial vector U_i.

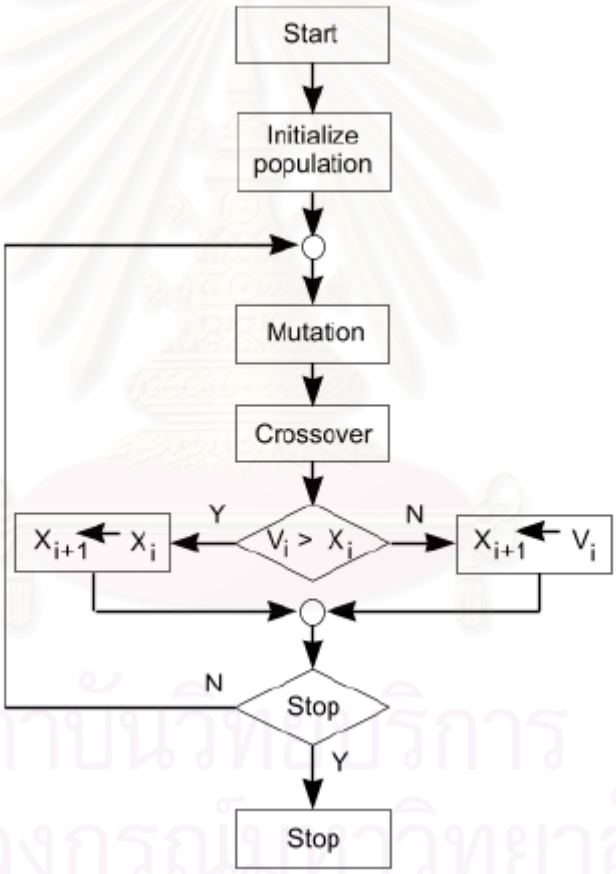


Figure 3.3 The flowchart of the differential evolutionary algorithm

The basic *DE/rand/1/bin* and trigonometric schemes are used in this algorithm. *DE/rand/1/bin* scheme starts by randomly selecting three vectors in the populations. The perturbed vector V_i is then generated based on the three previously selected vectors as follows:

$$V_i = X_{r3} + F(X_{r2} - X_{r1}) \tag{3.22}$$

where, X_{r1}, X_{r2} and X_{r3} are randomly selected vectors, and $r1 \neq r2 \neq r3 \neq i$ are satisfied. $F \in [0; 1+]$ is a control parameter of the algorithm. The trigonometric mutation scheme also starts by randomly selecting three vectors in the populations as in the *DE/rand/1/bin* scheme. But, the perturbed variable is calculated using the center point of the hyper geometric triangle of three previously selected vectors. The perturbed vector V_i is then generated by perturbing the center point a sum of three weighted vector differentials, as described by the following formulation:

$$\begin{aligned} V_i = {} & \frac{(X_{r1} + X_{r2} + X_{r3})}{3} + (p_2 - p_1)(X_{r1} - X_{r2}) \\ & + (p_3 - p_2)(X_{r2} - X_{r3}) + (p_1 - p_3)(X_{r3} - X_{r1}) \end{aligned} \tag{3.23}$$

where:

$$p_1 = \frac{|f(X_{r1})|}{|f(X_{r1})| + |f(X_{r2})| + |f(X_{r3})|}$$

$$p_2 = \frac{|f(X_{r2})|}{|f(X_{r1})| + |f(X_{r2})| + |f(X_{r3})|}$$

$$p_3 = \frac{|f(X_{r3})|}{|f(X_{r1})| + |f(X_{r2})| + |f(X_{r3})|}$$

Where, X_{r1}, X_{r2} and X_{r3} are randomly selected vectors, and $r1 \neq r2 \neq r3 \neq i$ are satisfied.

The perturbed vector $V_i(v_{i,1}, v_{i,2}, ..., v_{i,n})$ and its parent vector $X_i(x_{i,1}, x_{i,2}, ..., x_{i,n})$ are subjected to the crossover operation, which finally generates the trial vector $U_i(u_{i,1}, u_{i,2}, ..., u_{i,n})$ as follows:

$$u_{i,j} = \begin{cases} v_{i,j,} & if \quad random[0,1) \leq CR \vee j = random[1,n); \\ x_{i,j}, & otherwise \end{cases}$$

(3.24)

Where $CR \in [0, 1]$ is crossover factor. The created trial vector U_i is then compared with its parent vector X_i. If the trial vector is better than the parent vector, the trial vector replaces its parent vector in the population, as expressed in the following formulation:

$$X_{i+1} = \begin{cases} U_{i,} & if \quad f(U_i) \le f(X_i); \\ X_i, & otherwise \end{cases} \tag{3.25}$$

The evolutionary process repeats until the stopping criteria are satisfied.

3.5.2 The Constrain Handling Scheme

3.5.2.1 Handling Integer and Discrete Variables

The original DE is incapable of handling discrete variables. However, it is very easy to modify the algorithm to deal with integer and/or discrete variables. First, continuous variables are converted to integer variables by truncation. Then, the truncated variables are used to evaluate the objective function. It can be expressed using the following expression:

$$x' = (\text{int})x; \tag{3.26}$$

Discrete variables can also be easily handled. Instead of directly using discrete variables as the optimized variables, the index of all discrete variables is assigned first. The index of each discrete variable is then used as the optimized variables. But, to evaluate the objective function, the original discrete variables are used.

3.5.2.2 Handling Boundary Constraints

It is important that the optimize variables must lie inside their allowed ranges. We replace each variable that violates boundary constraints by the upper or lower limits, according to the following rule:

$$x' \begin{cases} x_i^{(L)}, & if \quad x < x_i^{(L)}; \\ x, & if \quad x_i^{(L)} \le x \le x_i^{(U)}; \\ x_i^{(U)}, & if \quad x > x_i^{(U)}; \end{cases} \tag{3.27}$$

Where, $x_i^{(L)}$ and $x_i^{(U)}$ are the upper and lower bounds of each variable, respectively.

3.5.2.3 Dominance-based Selection Scheme

A dominance-based selection scheme is used to incorporate constraints into the fitness function. When comparing trial vector U_i with its parent vector X_i, we can have three possible situations. In the first case, both U_i and X_i are feasible. The vector with a better objective function survives to the next generation. In the second case, one is feasible, but the other one is infeasible. The feasible vector survives to the next generation. In the last case, where both vectors are infeasible. The vector with lower degree of constraints violation survives to the next generation. The rule for the selection is defined as follows:

$$X_{i+1} = \begin{cases} X_{i,} & X_i \prec U_i; \\ U_i, & otherwise; \end{cases} \tag{3.28}$$

Where, $X_i \prec U_i$ denotes that X_i dominates U_i. That is X_i has better objective function than U_i and/or lower degree of constraints violation.

3.5.2.4 Handling Equality Constraints

Generally, the equality constraints can be used to reduce the number of dimensions for the optimization problem without distorting the results. However, identifying the reduced variables is still a hard task. Moreover, some equality constraints are irreducible, and cannot be used to transform the problem to the lower dimension problem. Consider the case of n-dimensional optimization problem with *m* equality constraint *(H(X) = 0)*, the degree of freedom for this problem is actually *n - m*. That is only *n-m* variables are independent, while *m* variables are defined by the equality constraints. Therefore, any infeasible vector *X* containing *n* variables can be repaired by solving the system of m equations. Newton's method herein is applied to solve the system of equality constraint equations. In the first step, *m* variables from

totally *n* variables are randomly selected to be repaired. The degree of constraints violation are checked whether it is greater than a specified tolerance *e*. Infeasible vectors with small degree of violation are allowed to survive. This helps to maintain diversity in the population. On the other hand, infeasible vectors with large degree of violation are then repaired by solving the system of m equations. The corrected vector X that is computed by equation 3.22 moves each equality constraint closer to the allowable range.

$$X_{i+1} = X_i - J^{-1}(X_i)H(X_i) \tag{3.29}$$

where, $J(X_i)$ is the Jacobian matrix, and $H(X_i)$ is the vector of equality constraints violation. Iteration stops if either the sum of the degree of constraints violation is less than a given tolerance e, or the maximum iteration number has been reached.

CHAPTER IV

NYLON 6 PRODUCTION PROCESS

4.1 Process Description

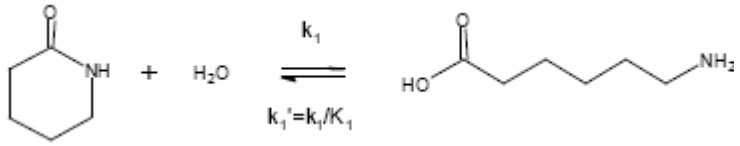
Nylon 6 or polycaprolactam is a polymer developed to reproduce the properties of nylon 6,6 without violating the patent on its production. Unlike most other nylons, nylon 6 is not a condensation polymer, but instead is formed by ring-opening polymerization. This makes it a special case in the comparison between condensation and addition polymers. Nylon 6 is widely used in the synthetic fiber industry, automotive parts, electrical cables, and packaging.

4.1.1 Theory Nylon 6 Polymerization Reaction

Nylon 6 reaction starting from caprolactam by using water as catalyst is not as simple as the reaction for generating polyolefines or polyester. The reaction are carried out in five steps:

Step 1: Ring Opening of Caprolactam (Start Reaction)

The reaction initiated by water(W) and generating amino caproic acid(P_1). Start reaction of ring opening of caprolactam(CL) is a slow, endothermic reaction, accelerated by water and temperature, catalyst by COOH-groups. The formula of the chemical reaction can be described as follows:



caprolactam water amino caproic acid

Figure 4.1 Ring opening of caprolactam via water

The functional group notation for this reaction is:

$$CL + W \underset{k_1' = k_1 / K_1}{\overset{k_1}{\rightleftharpoons}} P_1 \qquad (4.1)$$

Step 2: Polycondensation (Chain Growth)

This reaction step regulates the polymerization degree and herewith also the relevent properties of the polymer. The polymerization degree is proportional to the number of caprolactam molecules, which are built in the chain, and depends on the water content of the polymer. The formula of the chemical reaction can be described as follows:

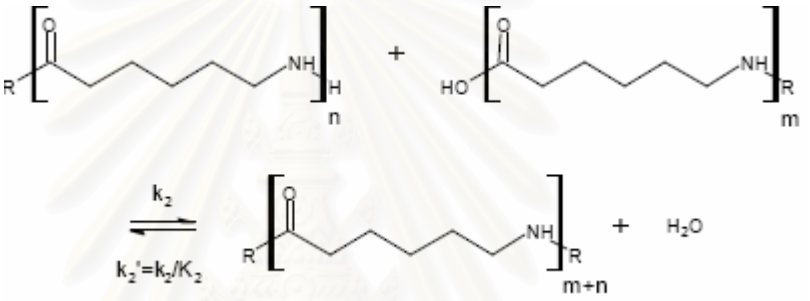


Figure 4.2 Polycondensation reaction

Two nylon-6 chains, of degree of polymerization n and m, react to form one longer chain of degree of polymerization n+m. Water is the small molecule that is eliminated. The group *R* can stand for hydrogen, a hydroxyl group, or a terminator group show in Figure 4.2

The polycondensation reaction inherently contains many possibilities for reaction between polymeric species. In order to enumerate these possibilities, we must consider the attack of any polymeric species with an amine end group on any polymeric species with a carboxylic end group. P_1 and terminal amine groups (T-NH_2) on polymer chains can attack the carboxyl groups on P_1 and terminal carboxyl groups (T-COOH) on polymer chains. We show all of these reactions below. (Kevin, 2003)

$$P_1 + P_1 \underset{k_2' = k_2 / K_2}{\overset{k_2}{\rightleftharpoons}} T - COOH : T - NH_2 + W \qquad (4.2)$$

$$P_1 + T - COOH \underset{k_2' = k_2 / K_2}{\overset{k_2}{\rightleftharpoons}} T - COOH : B - ACA + W \qquad (4.3)$$

$$T - NH_2 + P_1 \underset{k_2' = k_2 / K_2}{\overset{k_2}{\rightleftharpoons}} T - NH_2 : B - ACA + W \qquad (4.4)$$

$$T-NH_2+T-COOH \underset{k_2^{'}=k_2/K_2}{\overset{k_2}{\rightleftharpoons}} B-ACA:B-ACA+W \qquad (4.5)$$

Step 3: Polyaddition of Caprolactam (Caprolactam Conversion)

The reaction between caprolactam and already generated amino caproic acid. The polyaddition is an exothermic reaction and is carried out faster than the polycondensation reaction. Therefore this reaction influences strongly the caprolactam turnover. The formula of the chemical reaction can be described as follows:

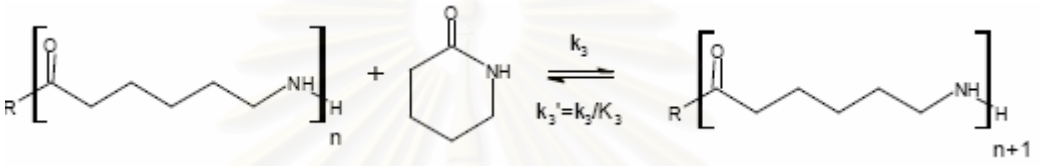


Figure 4.3 Addition of caprolactam

Any polymer with amine functionality can perform the forward reaction. We list all of the possibilities below:

$$P_1+CL \underset{k_3^{'}=k_3/K_3}{\overset{k_3}{\rightleftharpoons}} T-NH_2:T-COOH \qquad (4.6)$$

$$T-NH_2+CL \underset{k_3^{'}=k_3/K_3}{\overset{k_3}{\rightleftharpoons}} T-NH_2:B-ACA \qquad (4.7)$$

Step 4: Ring Opening of Cyclic Dimer

Cyclic dimer can be opened by water; this reaction is analogous to ring opening of caprolactam. The formula of the chemical reaction can be described as follows:

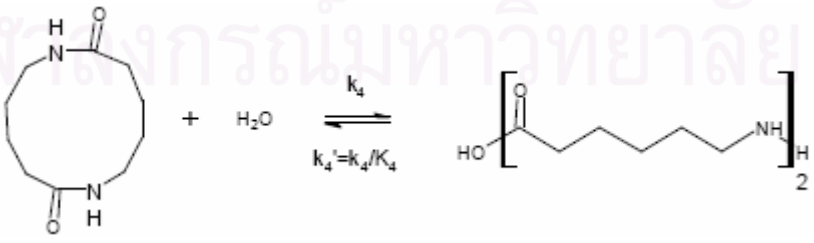


Figure 4.4 Ring opening of cyclic dimer

The functional group notation for this reaction is:

$$CD + W \underset{k_4^{'}=k_4/K_4}{\overset{k_4}{\rightleftharpoons}} T-COOH : T-NH_2 \qquad (4.8)$$

Step 5: Polyaddition of Cyclic Dimer

We also consider polyaddition of cyclic dimer; this reaction is analogous to the polyaddition of caprolactam. The formula of the chemical reaction can be described as follow:

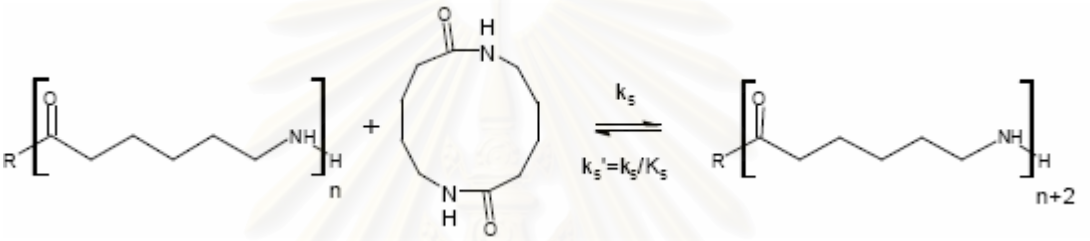


Figure 4.5 Polyaddition of cyclic dimer

A terminal amine group of any polymer can perform this addition. Therefore, over all polymer species, we have:

$$P_1 + CD \underset{k_5^{'}=k_5/K_5}{\overset{k_5}{\rightleftharpoons}} T-NH_2 : B-ACA : T-COOH \qquad (4.9)$$

$$T-NH_2 + CD \underset{k_5^{'}=k_5/K_5}{\overset{k_5}{\rightleftharpoons}} B-ACA : B-ACA : T-NH_2 \qquad (4.10)$$

The rate constant for 5 equilibrium reactions. We can follow as (Kevin, Neeraj & Liu, 2003) in Table 4.1

Table 4.1 Rate Constants for the Equilibrium Reactions in equation 4.1 – 4.10

rate constant expression	$k_i = A_i^0 \exp\left(-\frac{E_i^0}{RT}\right) + A_i^c \exp\left(-\frac{E_i^c}{RT}\right)[T - COOH]$
equilibrium constant expression	$K_i = \frac{k_i}{k_i'} = \exp\left(\frac{\Delta S_i - \Delta H_i / T}{R}\right)$

i	A_i^0 (kg/mol*s)	E_i^0 (J/mol)	A_i^c (kg^2/mol^2*s)	E_i^c (J/mol)	ΔH_i (J/mol)	ΔS_i J/mol*K
1	1.66E+02	8.32E+04	1.20E+04	7.87E+04	8.03E+03	-33.01
2	5.26E+06	9.74E+04	3.37E+06	8.65E+04	-2.49E+04	3.951
3	7.93E+05	9.56E+04	4.55E+06	8.42E+04	-1.69E+04	-29.08
4	2.38E+08	1.76E+05	6.47E+08	1.57E+05	-4.02E+04	-60.79
5	7.14E+04	8.92E+04	8.36E+05	8.54E+04	-1.33E+04	2.439

4.1.2 General Process description

4.1.2.1 Caprolactam feeding part

In this section, caprolactam bags are prepared for the folowing melting in the melter, for the use in the polymerization process. The delivered caprolactam bags are manully opened by a knive and emptied in funnel device. Lactam is then crushed and fed to the lactam melter.

4.1.2.2 Monomer part

In this section, caprolactam is prepared for use in the polymerization process. Solid caprolactam is charged to the melter and melted under nitrogen atmosphere by a hot water jacket and an external steam heated heat exchanger. The lactam is circulated through filters back to the melter. The molten caprolactam is then stored in a lactam tank for further metering to the polymerization the lactam section. Before entering the pressure polymerization the lactam is mixed in a fixed ratio with recovered lactam, filtered and heated up in a pre-heater.

4.1.2.3 Additive System

Condensate or demineralized water used as intiator is filtered and fed by gravity into receiver is dosed continuously into the lactam stream by metering pump.

4.1.2.4 Pressure Polymerizer

In the pressure stage mainly the ring opening reaction of the caprolactam is carried out and also a degree of polyaddition. This take place at elevated temperature and slight overpressure. The lactam stream entering the reactor is heated to reaction temperature in the upper section of the reactor. As the material moves downwards in the reactor, flow regulators ensure an uniform product flow during polymerization, to reach the required product viscosity.

The excess catalyst-water which is evaporated from the top of the pressure polymerizer reactor passes through a packed column, while the remainder flows to the Distillation. Afterwards the pre-polymer is fed to the final polymerizer by means of a discharge gear pump.

4.1.2.5 Final Polymerizer

In the final polymerizer the surplus of water is taken off to allow the polymer chains to grow to the desired polymerization degree. The reaction of polymerization takes place under vacuum and elevated temperature. The water/lactam vapors leaving the top of the Final polymerizer enter a partial condenser for separation. The vacuum is generated by means of water jets. The product increases in viscosity while flowing downwards through the reactor by a uniform flow ensured by built in flow distributors regulate not only the downwards stream but also the upwards stream of water bubbles. The product is then discharged to the die head by means of a gear pump.

Three heating system are provided for the heating of the Final polymerizer. The built in heat exchanger moderator, serves to remove the polymerization heat from the product, in order to prevent it from overheating. All heating system are heated electrically.

Schematic of industrial nylon 6 production process. (illustrated as Figure 4.6)

Most of the monomer conversion takes place in the first reactor, while the molecular weight build occurs in the second reactor.

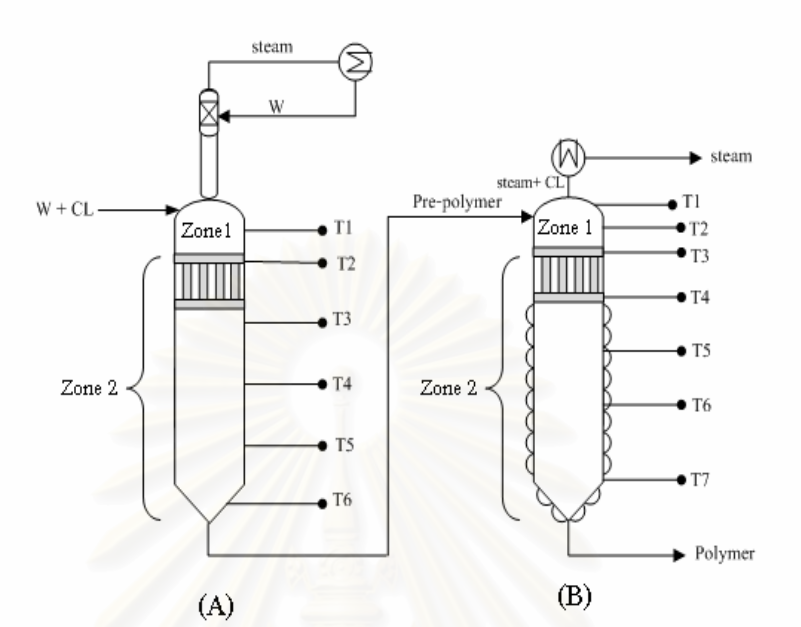


Figure 4.6 Two VK column reactors for industrial nylon 6 production process: (A) pressure polymerizer (B) final polymerizer

In this work, we apply data reconciliation in VK column reactors for industrial nylon 6 production process at steady state condition. And we study the performance of each available technique that suitable for industrial nylon 6 production process. Performance is compared between Robust function method and Weighted Least-Square. Robust function interested: Contaminated Normal, Lorentzian distribution function and Hampel's redescending M-estimator. Before, we study the performance of each available technique that suitable for industrial nylon 6 production process. We taking data reconciliation with gross error detection apply to other cases for performance comparison each algorithm. There are three cases:

- Application of data reconciliation with gross error detection in simple case
- Application of data reconciliation with gross error detection in nylon 6 production process by simulated at steady state condition
- Application of data reconciliation with gross error detection in industrial nylon 6 production process at steady state condition

In application of data reconciliation in three cases to compare performance. We divide two parts of system: system with all measured variables and system with all unmeasured variables. To efficient test of data reconciliation to use estimate the current status of process variables and unmeasured variables in all case could obtain efficient system.

4.2 Case Studies

4.2.1 Systems with All Measured Variables

4.2.1.1 Case A: Application of Data Reconciliation with Gross Error Detection in Simple Case

Let us first the simplest data reconciliation problem: the reconciliation of the stream flow of process. Initially, all flow rates are assumed to be directly measured. We assume a process operating at steady state condition.

In this case, we assume measured flow rates in process have 3 variables: A, B, C as show in Figure 4.7. Let us also ignore the energy flows of this process and focus only on the mass flows. We denote the true value of each variable are 1, 2 and 3 respectively.

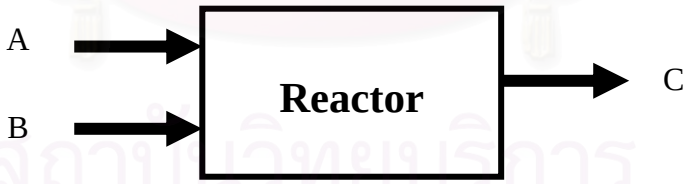


Figure 4.7 Example of process in simple case

The flow balance around the reactor can be written as:

$$A + B - C = 0 \tag{4.11}$$

In step compare the performances of each available technique that suitable for usability. We can classify into 3 cases.

- Case 1 : measurement data have not gross error

- Case 2 : measurement data have gross error
- Case 3 : measurement data contain both Normal and Uniform distributions

The measured values in case 1 – case 3 do not satisfy in equation 4.11. Therefore, we are desired to derive estimates of the flow that satisfy the above flow balance. The aim of reconciliation is to make minor adjustments to the measurements in order to make them consistent with the material balance. The adjusted measurements, which are referred to as estimates, are expected to be more accurate than the measurement. Therefore, we take data reconciliation with gross error detection apply this process and find algorithm that suitable for usability. There are step for performance test of each algorithms below:

Case 1: measurement data have not gross error

1. To determine measured data have normal distribution. While most of measurement values are distributed over true values range of each variables measured and Standard Deviation (SD) of distribution is 0.1. Therefore, measurement sets created can assume that there are only random error and lack of gross error.

2. Various different data reconciliation with gross error detection techniques are performed for find true solution of measurement data set each variables in process. The algorithms that we studied are as follows: Weighted Least-Square, Contaminated Normal**,** Lorentzian distribution function and Hampel's redescending M-estimator.

3. To compare performance each algorithms take solution to obtain each algorithm compare to true value of each variable at steady state condition.

Case 2: measurement data have gross error

1. To determine measured data have gross error. Take normal distribution measurement data added uniform distribution data in ratio 10%, 20% and 30% from old data. Therefore, measurement data created can assume that there are both random error and gross error 10%, 20% and 30% respectively.

2. Various different data reconciliation with gross error detection techniques are performed for find true solution of measurement data set each variables in process. The algorithms that we studied are as follows: Weighted Least-Square, Contaminated Normal**,** Lorentzian distribution function and Hampel's redescending M-estimator.

3. To compare performance each algorithms take solution to obtain each algorithm compare to true value of each variable at steady state condition.

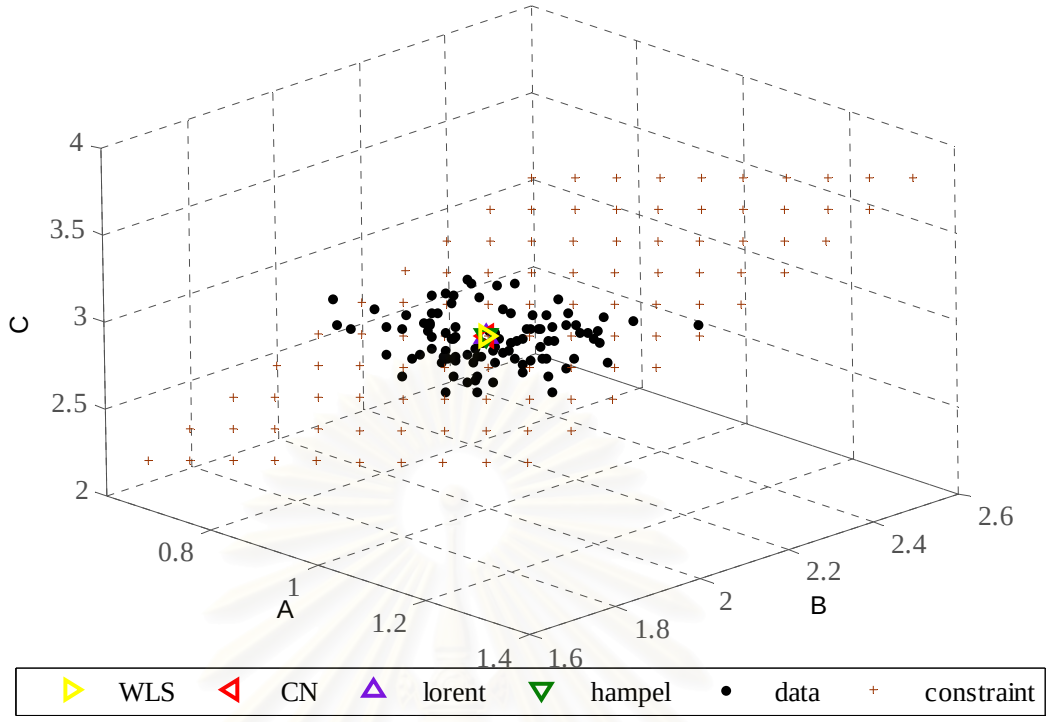
Case 3: measurement data contain both Normal and Uniform distributions

1. Using true values such as design data, measurement sets are created for each variable by adding noise from Normal and Uniform distributions with equal probability, i.e. half of the simulated measurement errors has a Normal probability distribution and the other half are from Uniform probability distribution.

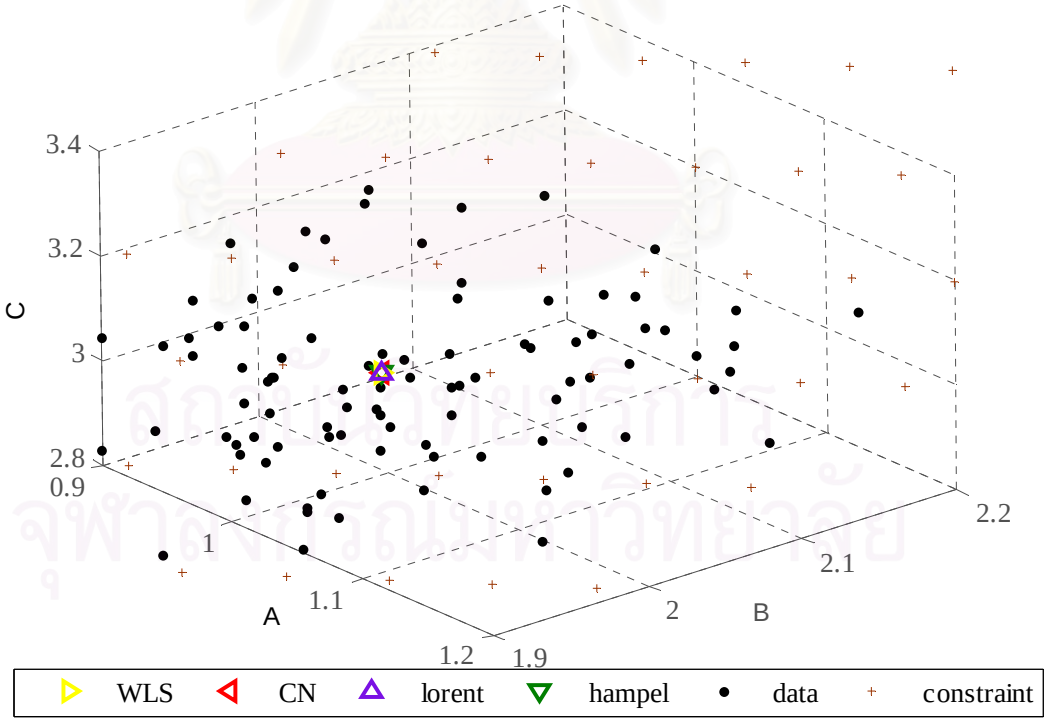
2. Various different data reconciliation with gross error detection techniques are performed for find true solution of measurement data set each variables in process. The algorithms that we studied are as follows: Weighted Least-Square, Contaminated Normal**,** Lorentzian distribution function and Hampel's redescending M-estimator.

3. To compare performance each algorithms take solution to obtain each algorithm compare to true value of each variable at steady state condition.

Reconciliation solution for the measured variables in case 1 through case 3 (shown in Figure 4.8 – Figure 4.13)

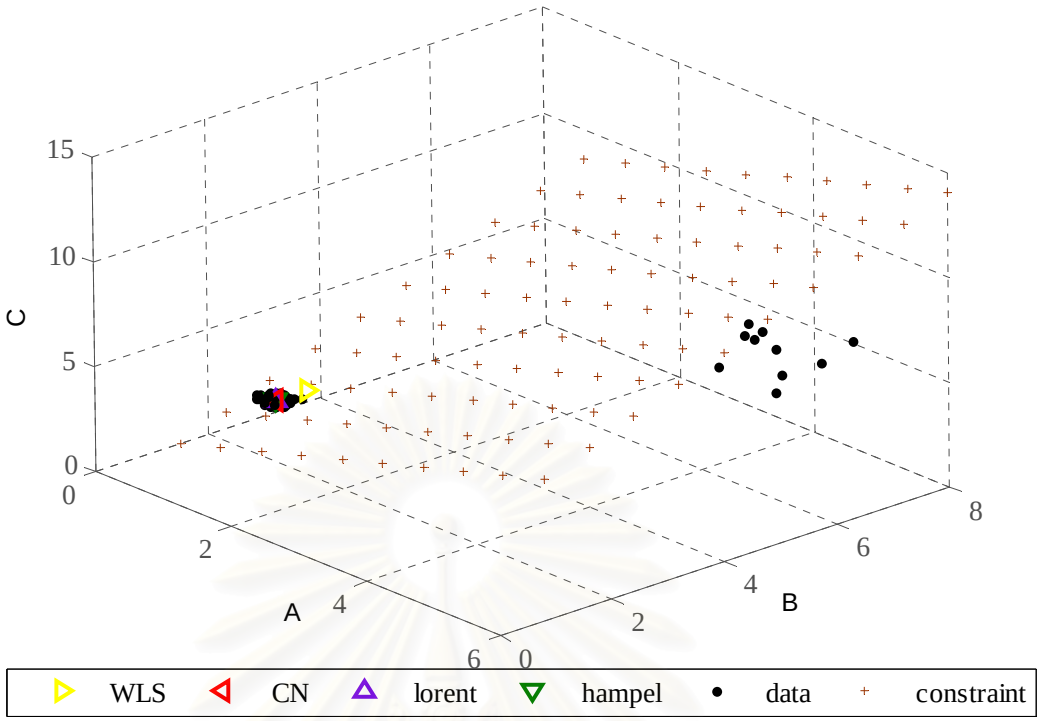


(a)

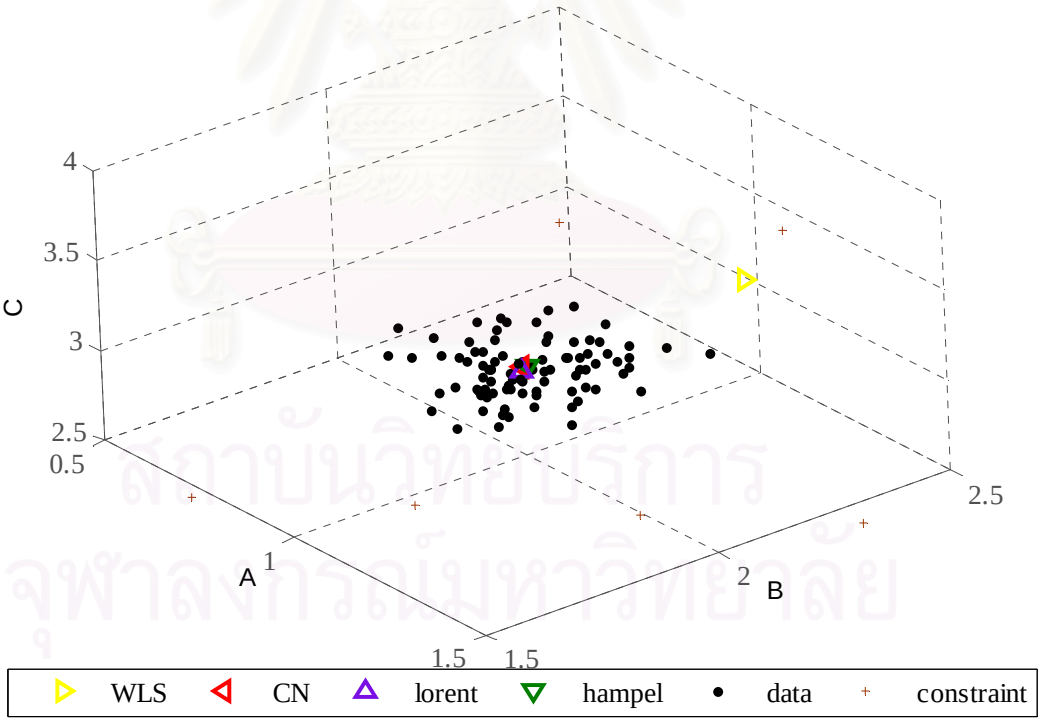


(b)

Figure 4.8 Distribution of measured A, B, C at have only random error:
(a) normal view (b) expansion view

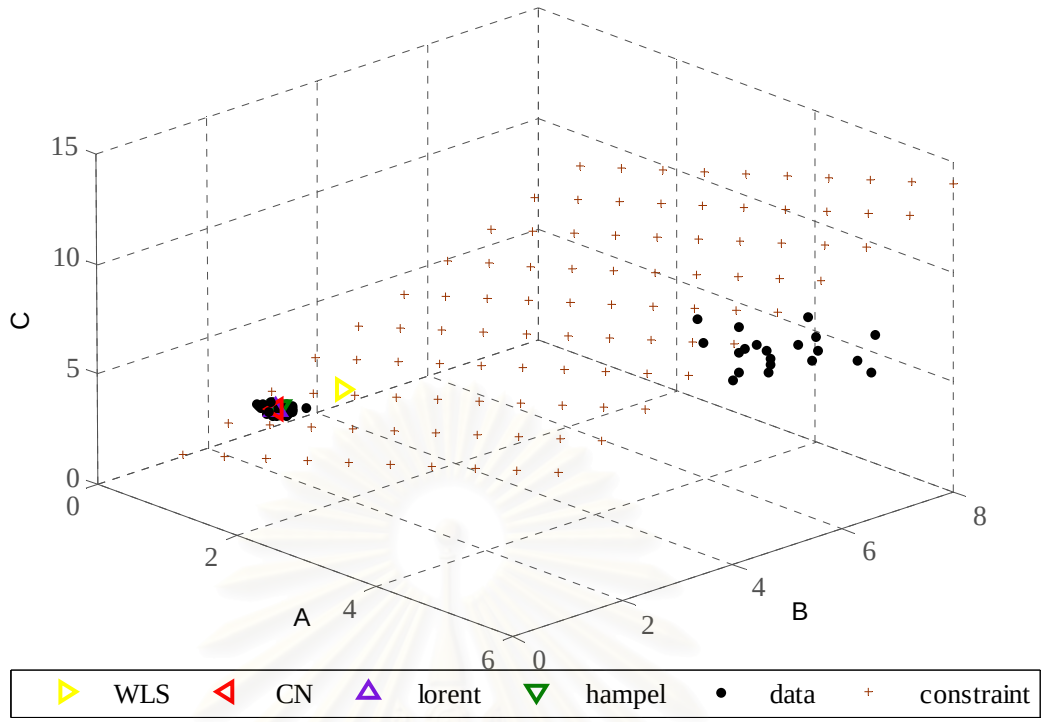


(a)

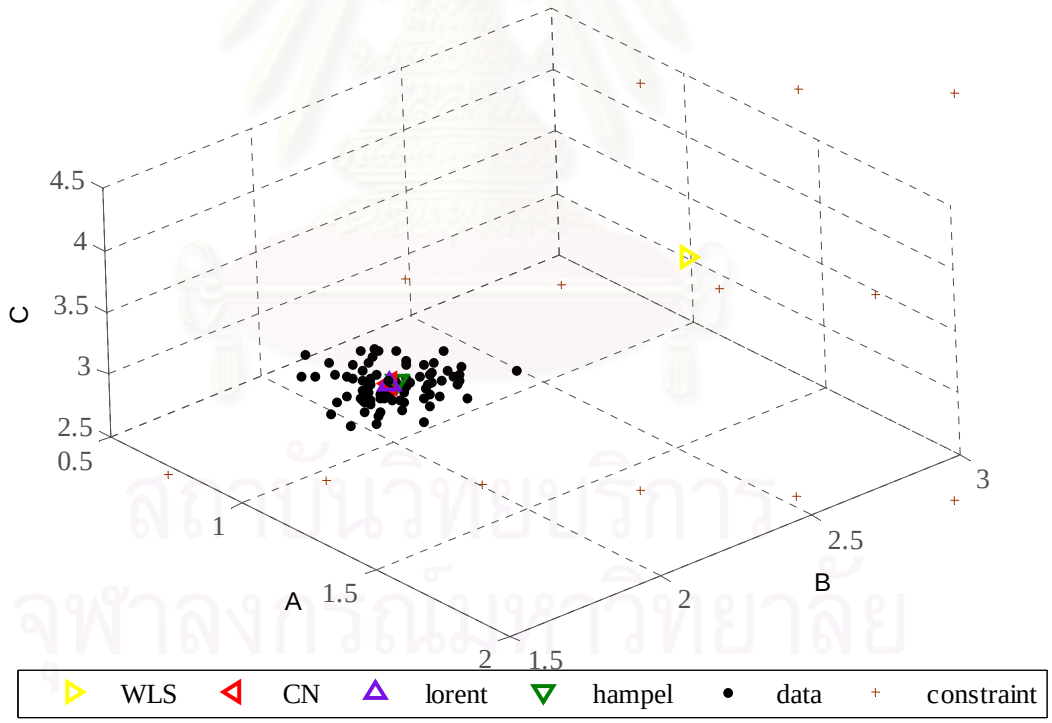


(b)

Figure 4.9 Distribution of measured A, B, C at have gross error 10%:
(a) normal view (b) expansion view

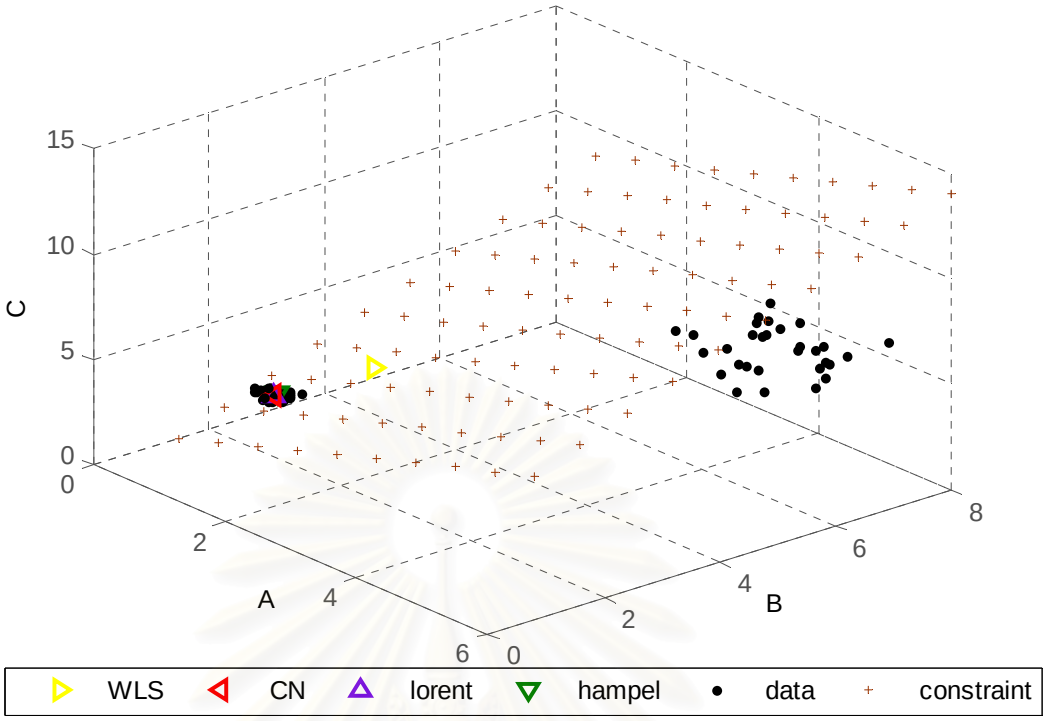


(a)

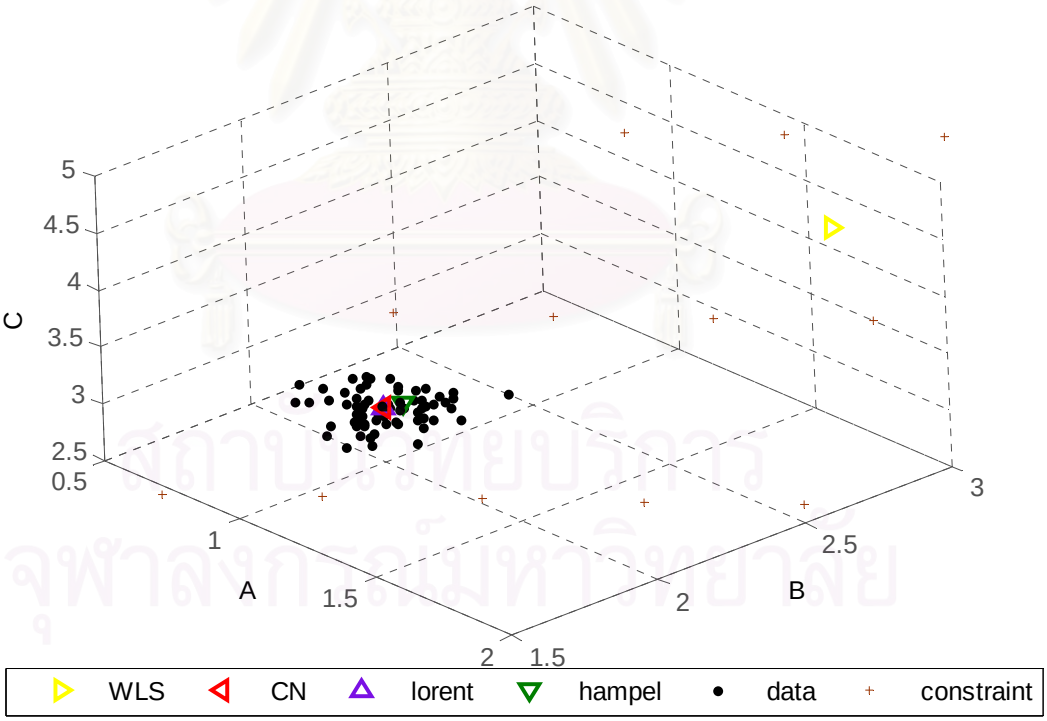


(b)

Figure 4.10 Distribution of measured A, B, C at have gross error 20%:
(a) normal view (b) expansion view

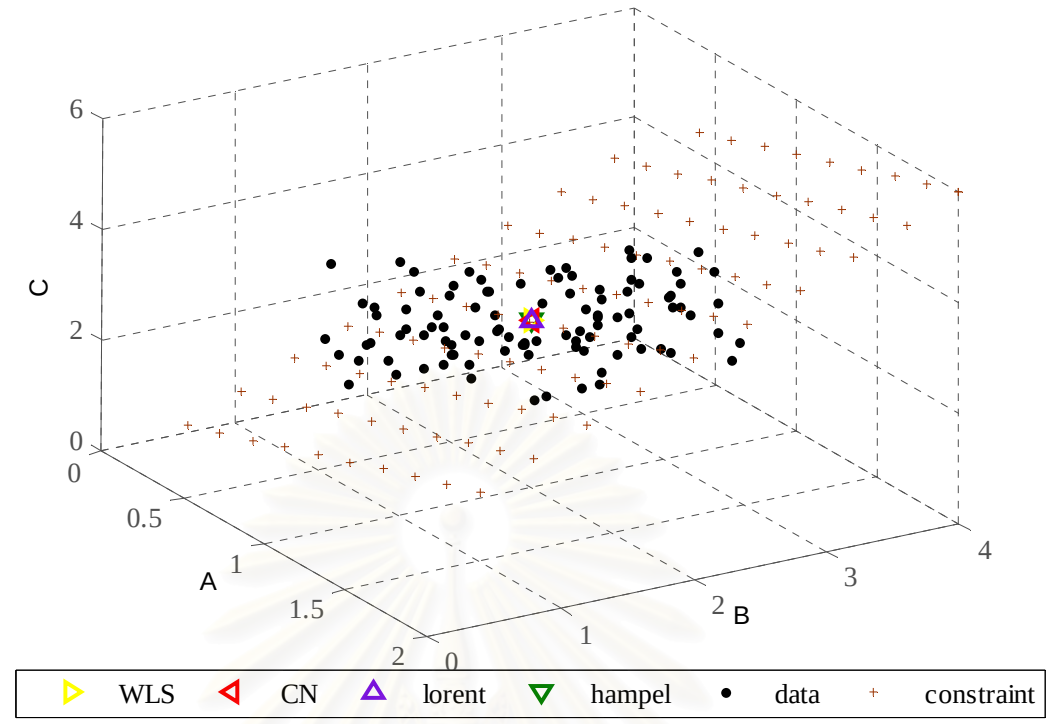


(a)

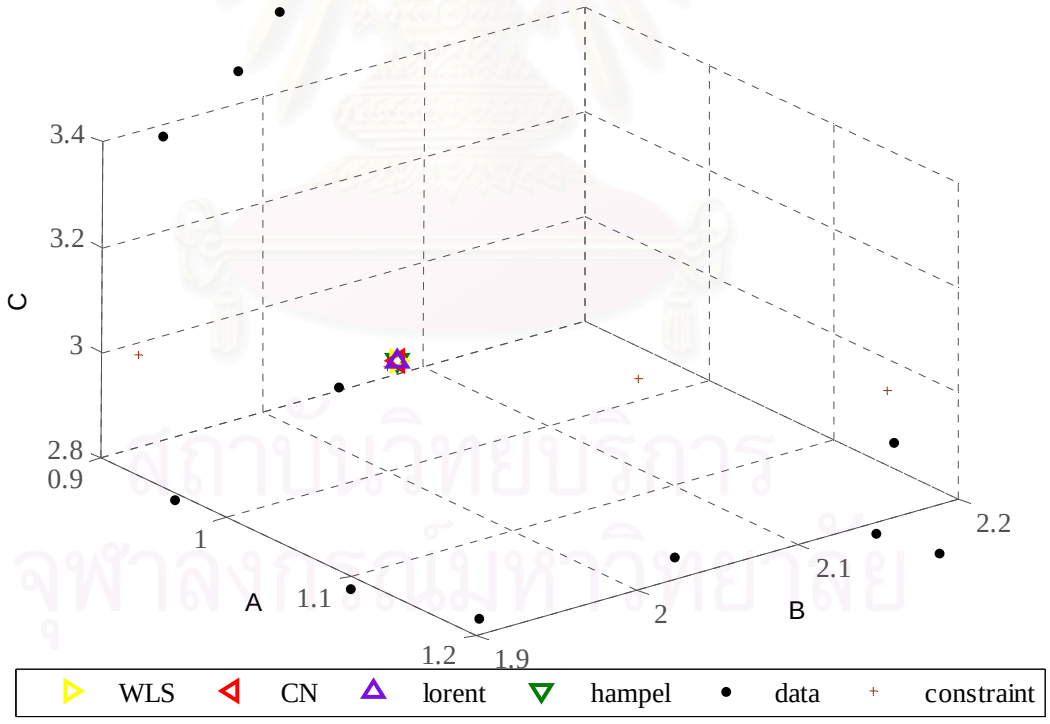


(b)

Figure 4.11 Distribution of measured A, B, C at have gross error 30%:
(a) normal view (b) expansion view



(a)



(b)

Figure 4.12 Distribution of measured A, B, C at contain both Normal and Uniform distributions: (a) normal view (b) expansion view

4.2.1.2 Case B: Application of Data Reconciliation with Gross Error Detection in Nylon 6 Production Process by Simulated at Steady State Condition

We simulate nylon 6 production process in Continuous stirred tank reactor (CSTR). The CSTR operates at 513.15 K for 12 hours. And total mass flow rate inlet is 100 kg/hr: 99 kg/hr caprolactam and 1 kg/hr of water. The reactor is operated at high pressure, i.e., we simulate it as a single liquid phase.

The assumptions are used in simulation as follows:

1. Concentration and temperature are distributed uniformly both in the reactor.
2. Total mass flow rate inlet equal total mass flow rate outlet
3. Simulate at steady state condition.

The material balance of each component (W, CL, CD, P_1, B-ACA, T-NH_2, T-COOH) in nylon 6 production process can be written as:

$$0 = \left(C_{W,in} - C_W\right)\left(\frac{F_{in}}{M}\right) + \left(R_2 + R_3 + R_4 + R_5\right) - \left(R_1 + R_8\right) \quad (4.12)$$

$$0 = \left(C_{CL,in} - C_{CL}\right)\left(\frac{F_{in}}{M}\right) - \left(R_1 + R_6 + R_7\right) \quad (4.13)$$

$$0 = \left(C_{CD,in} - C_{CD}\right)\left(\frac{F_{in}}{M}\right) - \left(R_8 + R_9 + R_{10}\right) \quad (4.14)$$

$$0 = \left(C_{P_1,in} - C_{P_1}\right)\left(\frac{F_{in}}{M}\right) + R_1 - \left(2R_2 + R_3 + R_4 + R_6 + R_9\right) \quad (4.15)$$

$$0 = \left(C_{B-ACA,in} - C_{B-ACA}\right)\left(\frac{F_{in}}{M}\right) + \left(R_3 + R_4 + 2R_5 + R_7 + R_9 + 2R_{10}\right) \quad (4.16)$$

$$0 = \left(C_{T-NH_2,in} - C_{T-NH_2}\right)\left(\frac{F_{in}}{M}\right) + \left(R_2 + R_6 + R_8 + R_9\right) - R_5 \quad (4.17)$$

$$0 = \left(C_{T-COOH,in} - C_{T-COOH}\right)\left(\frac{F_{in}}{M}\right) + \left(R_2 + R_6 + R_8 + R_9\right) - R_5 \quad (4.18)$$

The reaction velocity of equations (4.12)–(4.18) is follow as APPENDIX A in Table A.1, C_W, C_{CL}, C_{CD}, C_{P_1}, C_{B-ACA}, C_{T-NH_2} and C_{T-COOH} are the concentrations for the outlet stream(mol/kg), F_{in} is the total mass flow rate for the inlet

stream(kg/hr), and M is the total mass in the reactor(kg). Inlet stream includes caprolactam (CL) and water (W), outlet stream includes water, caprolactam, cyclic dimer(CD) and nylon 6 (P_n). In this case, we are interest mass flow rate of each component in inlet/outlet streams. Therefore, the mass flow rate balance around the process can be written as:

$$F_{W,in} + F_{CL,in} - F_W - F_{CL} - F_{CD} - F_{P_n} = 0 \tag{4.19}$$

Let us also ignore the energy mass flow rates this process and focus only on the mass flows. We believe that simulation at steady state is true value of process designed. A result obtained from nylon 6 productions simulation at steady state condition was presented in Table 4.2.

Table 4.2 Results of the nylon 6 production process simulation at steady state condition each variables

elements	mass flow rate (kg/hr)	
	inlet	outlet
water	1.00	0.845419
caprolactam	99.00	23.3329
cyclic dimer	-	0.316816
nylon 6	-	75.50477

In step compare the performances of each available technique that suitable for nylon 6 production process. We can divide into 3 cases:

- Case 1 : measurement data have not gross error
- Case 2 : measurement data have gross error
- Case 3 : measurement data contain both Normal and Uniform distributions

The measured values in case 1 – case 3 do not satisfy in equation 4.19. Therefore, we are desired to derive estimates of the flow that satisfy the above flow balance. Therefore, we take data reconciliation with gross error detection apply this process and find algorithm that suitable for nylon 6 production. There are steps for performance test of each algorithm below:

Case 1: measurement data have not gross error

1. To determine measured data have normal distribution. While most of measurement values are distributed over true values range of each variables measured and Standard Deviation (SD) of distribution is 0.1. Therefore, measurement sets created can assume that there are only random error and lack of gross error.

2. Various different data reconciliation with gross error detection techniques are performed for find true solution of measurement data set each variables in process. The algorithms that we studied are as follows: Weighted Least-Square, Contaminated Normal**,** Lorentzian distribution function and Hampel's redescending M-estimator.

3. To compare performance each algorithms take solution to obtain each algorithm compare to true value of each variable at steady state condition.

Case 2: measurement data have gross error

1. To determine measured data have gross error. Take normal distribution measurement data added uniform distribution data in ratio 10%, 20% and 30% from old data. Therefore, measurement data created can assume that there are both random error and gross error 10%, 20% and 30% respectively.

2. Various different data reconciliation with gross error detection techniques are performed for find true solution of measurement data set each variables in process. The algorithms that we studied are as follows: Weighted Least-Square, Contaminated Normal**,** Lorentzian distribution function and Hampel's redescending M-estimator.

3. To compare performance each algorithms take solution to obtain each algorithm compare to true value of each variable at steady state condition.

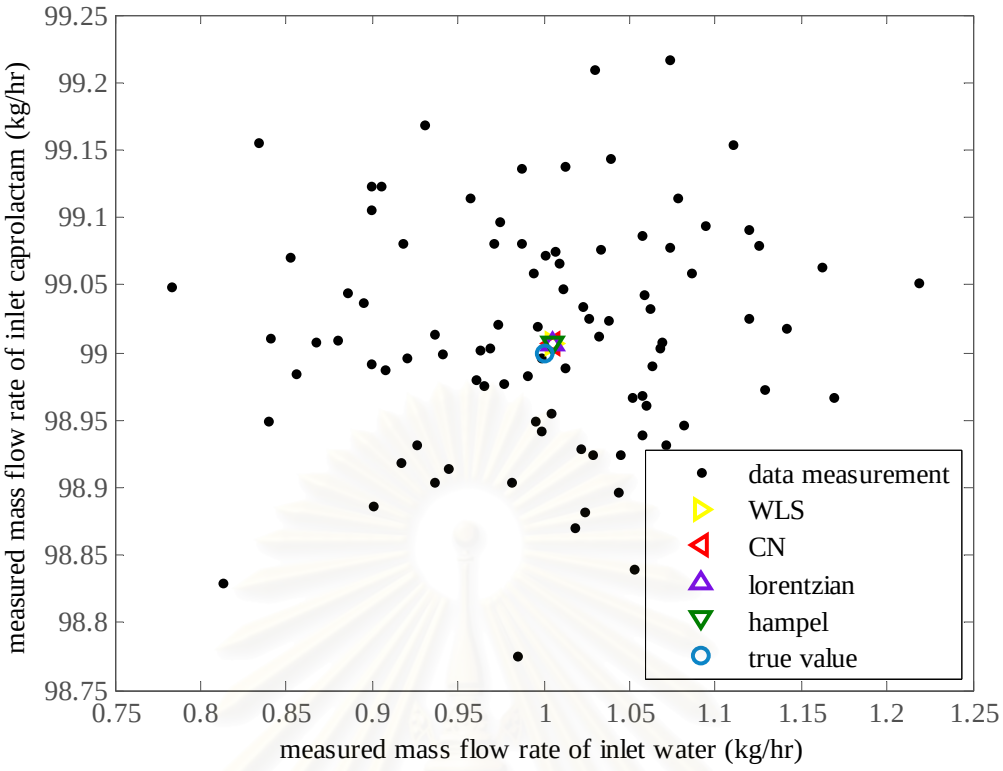
Case 3 measurement data contain both Normal and Uniform distributions

1. Using true values such as design data, measurement sets are created for each variable by adding noise from Normal and Uniform distributions with equal probability, i.e. half of the simulated measurement errors has a Normal probability distribution and the other half are from Uniform probability distribution.

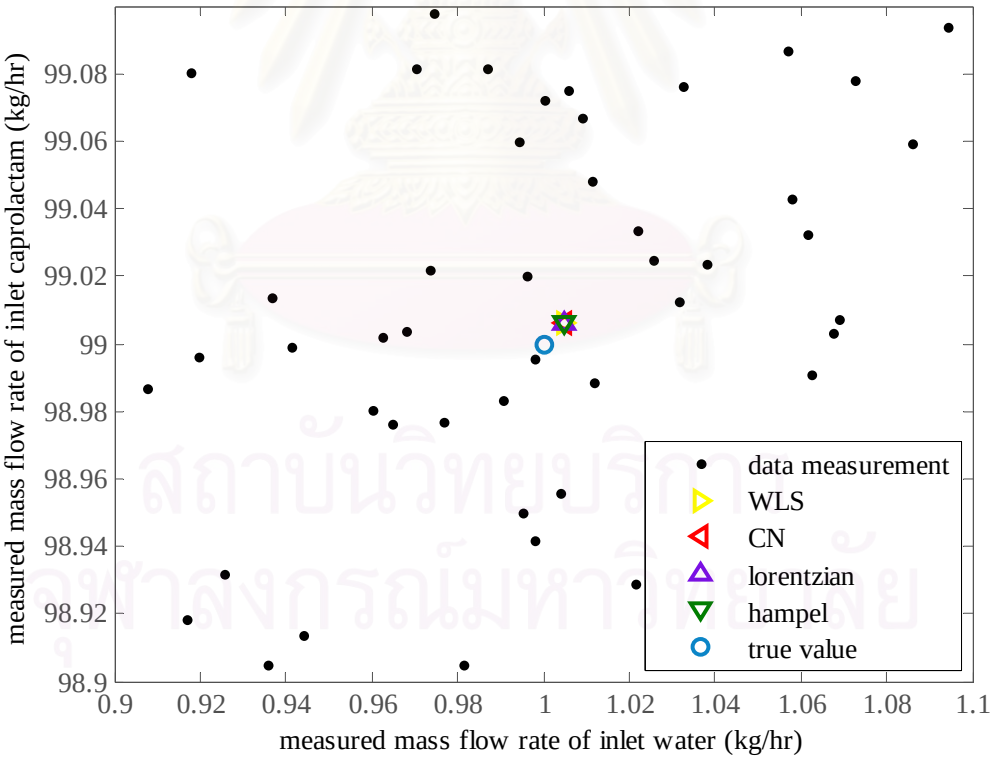
2. Various different data reconciliation with gross error detection techniques are performed for find true solution of measurement data set each variables in process. The algorithms that we studied are as follows: Weighted Least-Square, Contaminated Normal**,** Lorentzian distribution function and Hampel's redescending M-estimator.

3. To compare performance each algorithms take solution to obtain each algorithm compare to true value of each variable at steady state condition.

In this case, we are select variable for showing result of data reconciliation. There are analyzed 2 variables: mass flow rate of inlet water and caprolactam. Reconciliation solution for the measured variables in case 1 through case 3 (Figure 4.13 – Figure 4.17)

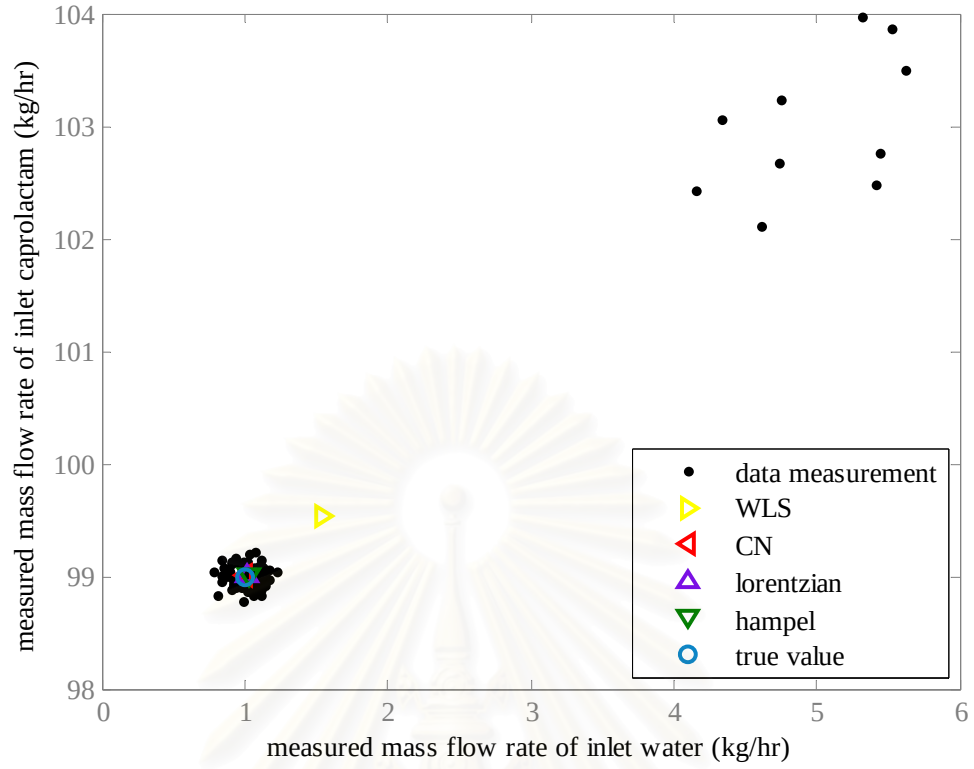


(a)

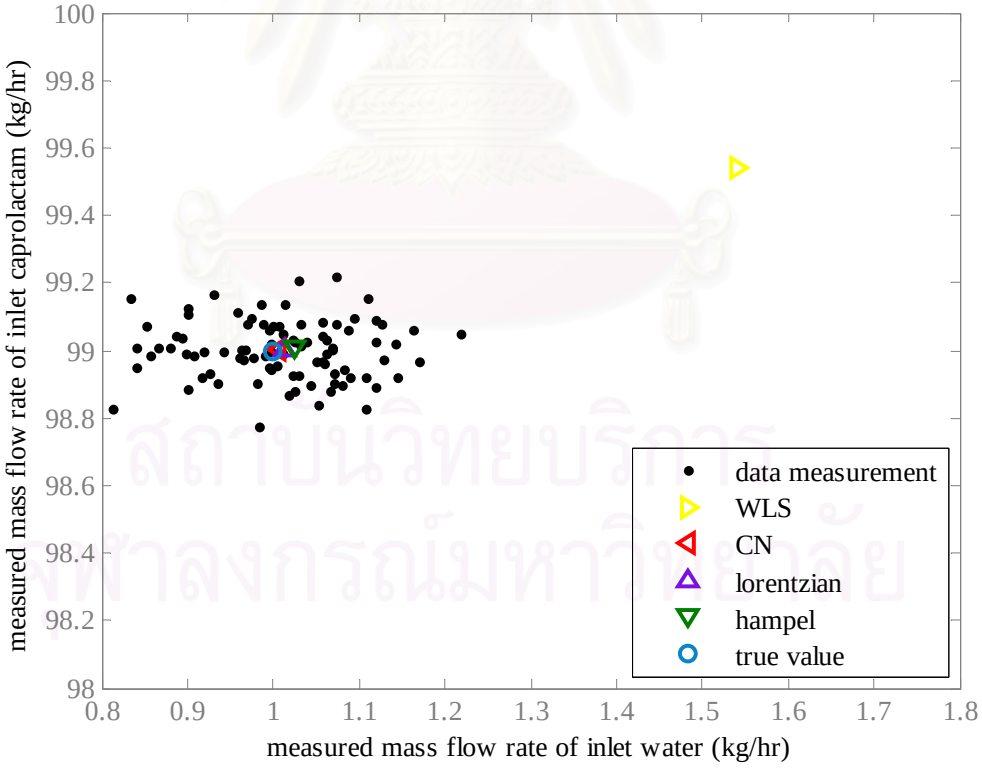


(b)

Figure 4.13 Distribution of measured mass flow rate of inlet water and caprolactam at have only random error (a) normal view (b) expansion view

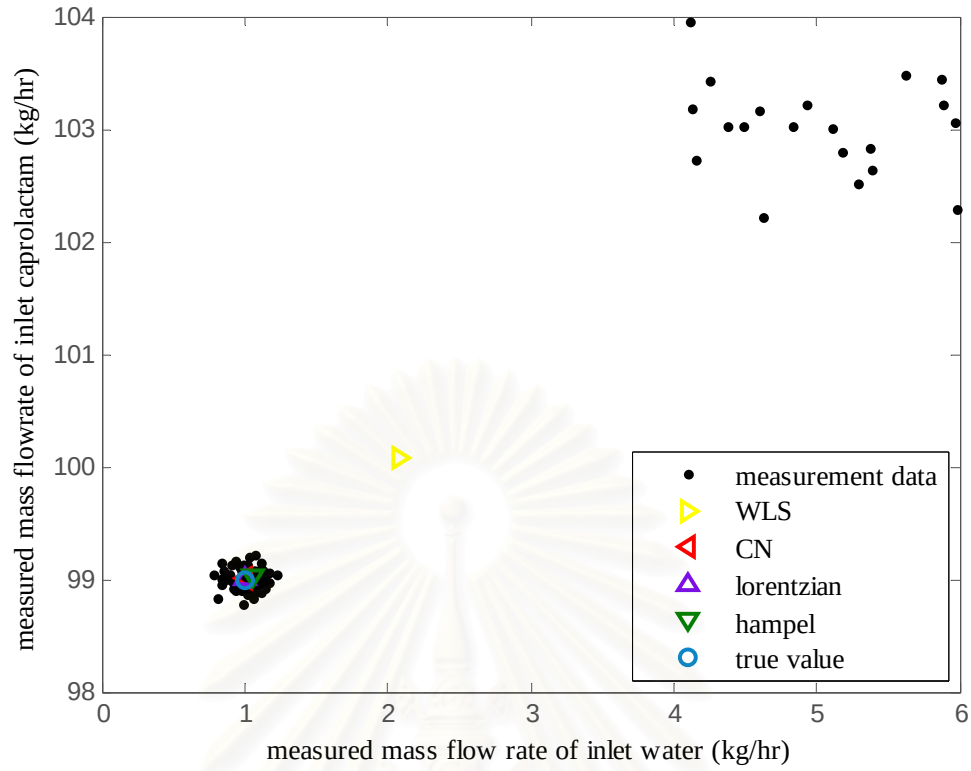


(a)

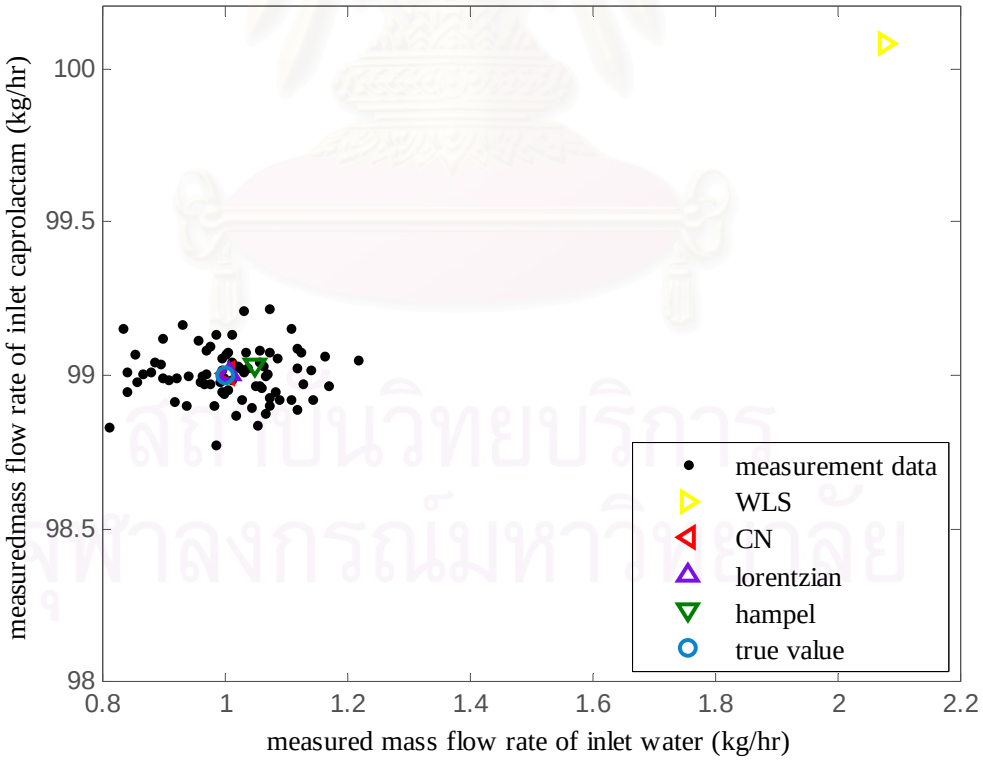


(b)

Figure 4.14 Distribution of measured mass flow rate of inlet water and caprolactam at have gross error 10% (a) normal view (b) expansion view

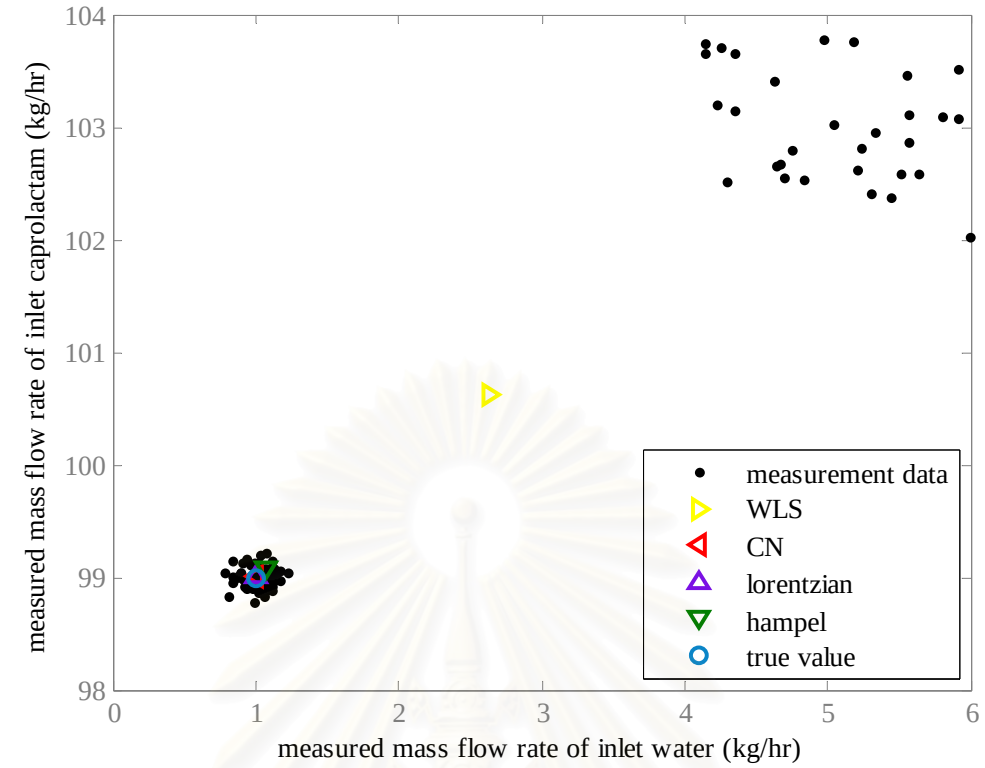


(a)

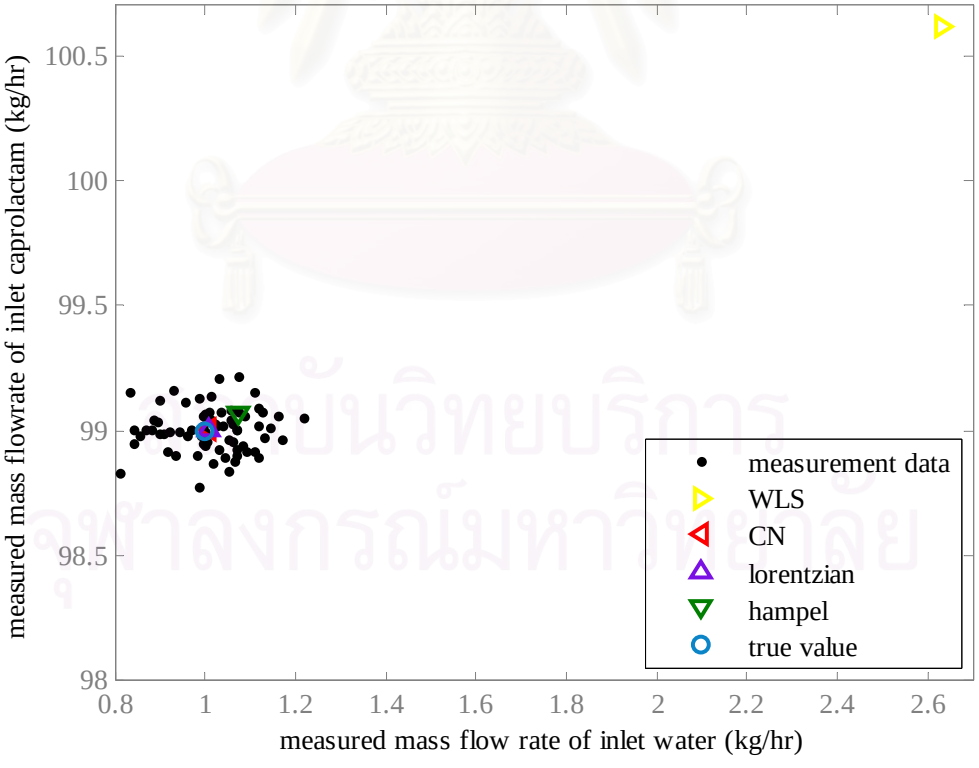


(b)

Figure 4.15 Distribution of measured mass flow rate of inlet water and caprolactam at have gross error 20% (a) normal view (b) expansion view

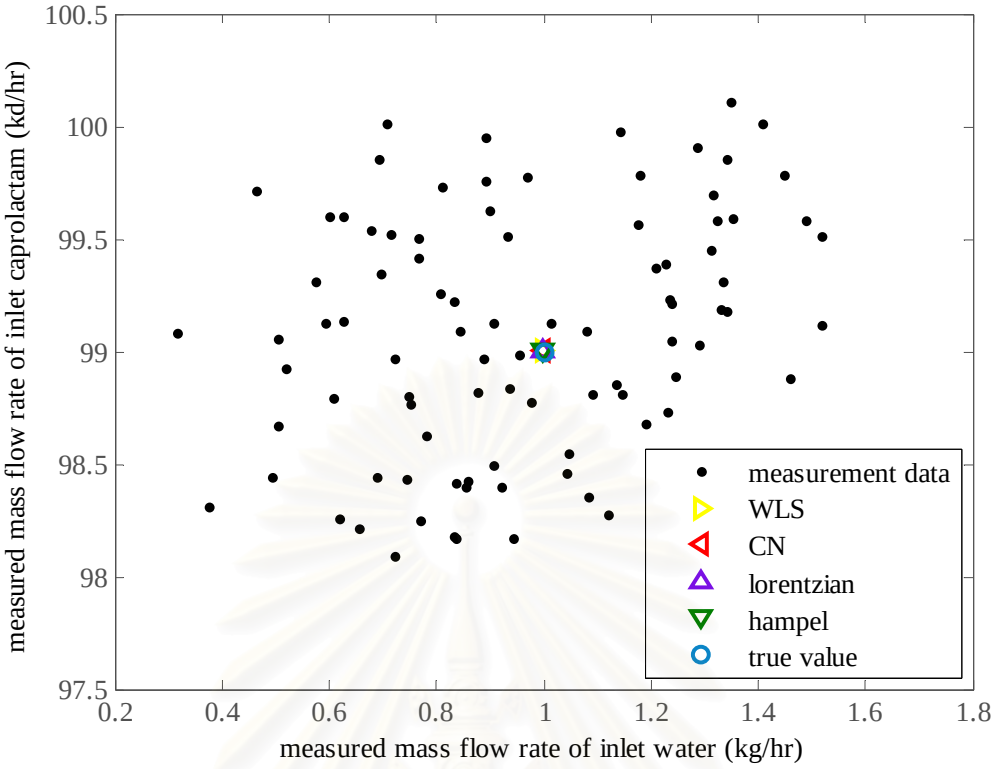


(a)

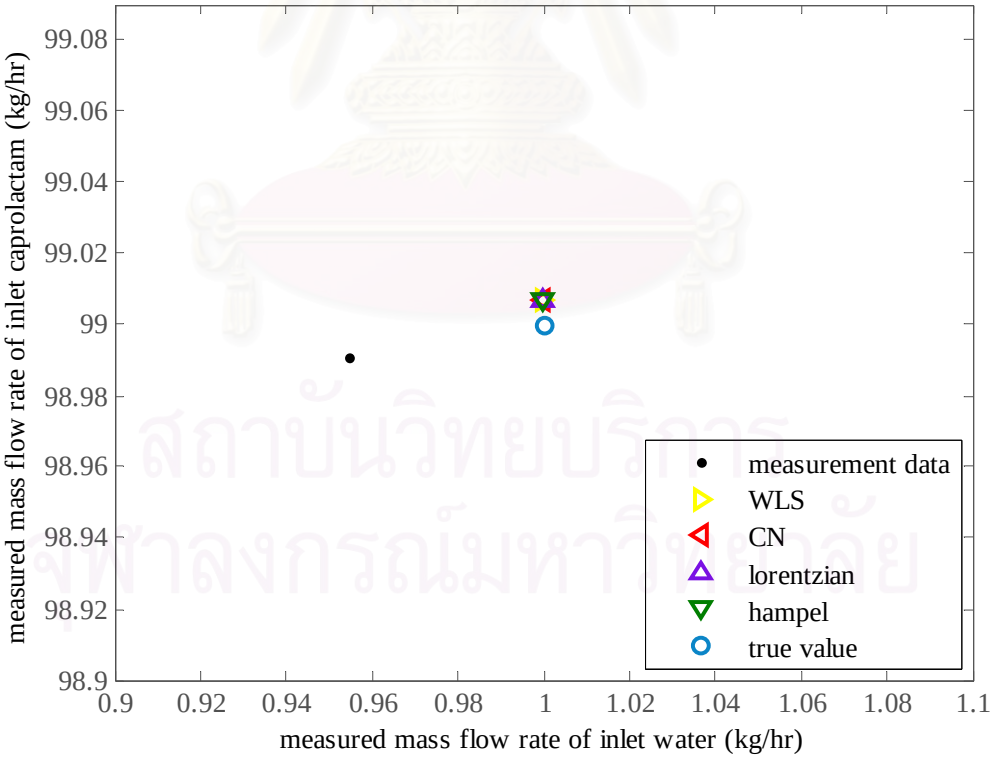


(b)

Figure 4.16 Distribution of measured mass flow rate of inlet water and caprolactam at have gross error 30% (a) normal view (b) expansion view



(a)



(b)

Figure 4.17 Distribution of measured mass flow rate of inlet water and caprolactam at contain both Normal and uniform distributions (a) normal view (b) expansion view

As a result of performance test both 2 cases are application of data reconciliation with gross error detection in simple case and nylon 6 production process by simulated at steady state condition. As an example in case 1, the measured data have normal distribution. The solutions obtained from apply data reconciliation each technique shown in Figure 4.8 and Figure 4.13.

As can be seen from Figure 4.8 and Figure 4.13. The result shows that in the case of process measurements contain only random error. The reconciled data obtained from 4 algorithms are close to true value at steady state condition. One of the possible reasons is that measurement data of each variables in process have normal distribution. While most of measurement values are distributed over mean range of each measured variables. Therefore, all algorithms perform well in this case.

For comparison, in the case that process measurements contain both gross and random errors, the measured data have gross errors present in measurements 10%, 20% and 30% respectively. The solution obtained from apply data reconciliation each techniques shown in Figure 4.9 and Figure 4.14 for gross error 10%, Figure 4.10 and Figure 4.15 for gross error 20% and Figure 4.11 and Figure 4.16 for gross error 30%.

As a result of reconciled data each algorithms in case 2. The finding demonstrated that reconciliation solution obtained from Weighted Least-Square method has gross errors present in measurements 10%, 20% and 30%. It incorrect to true value of process at steady state have stronger tendency to be increase with an increase amount of gross error. On the other hand, Contaminated Normal, Lorentzian distribution function and Hampel's redescending M-estimator approaches show a better performance. Due to the objective functions of Contaminated Normal distribution function, Lorentzian distribution function and Hampel's redescending M-estimator approaches are formed using the probability distribution function of the measured variable, by maximizing the product of individual probability values for each measured variables. Therefore, with 10%, 20% or 30% of gross errors in measurement does not affect to reconciled data obtained. However, the concept of Weighted Least-Square is to minimize difference between true value and measurement value. In the case of measurement data have gross error, the gross error

are also used in reconciliation. Therefore, reconciled data obtained from Weighted Least-Square are different to true value of each variable.

For comparison, the measured data contain both Normal and Uniform distributions. The solutions obtained from apply data reconciliation each technique shown in Figure 4.12 and Figure 4.17.

As a result of the reconciled data obtained from 4 algorithms are close to true value at steady state condition. It can also be said that all 4 methods give the same performance in case of measured data has pattern distribution. Nevertheless, if measured data have gross errors present in measurements. Weighted Least-Square method has generally given limited information concerning usability, because it can not support with gross errors present in measurements. Therefore, algorithms are the appropriate in case A and case B: Contaminated Normal distribution function, Lorentzian distribution function and Hampel's redescending M-estimator approaches. Because it can support with process measurements contain both gross and random errors.

The solutions obtained from apply data reconciliation each technique in all cases both case A and case B. We have shown in APPENDIX B.

4.2.1.3 Case C: Application of Data Reconciliation with Gross Error Detection in Industrial Nylon 6 Production Process at Steady State Condition

In this case, we apply data reconciliation with gross error detection in VK column reactors for industrial nylon 6 production process. Schematic of industrial nylon 6 production process (illustrated as Figure 4.6). As can be seen from Figure 4.6 the reactor for industrial nylon-6 production process has 2 reactors: (A) pressure polymerizer (B) final polymerizer. Therefore, the data reconciliation applied both 2 reactors. The motivation for reconciling these measurements arises from the need to estimate true value of measured mass flow rate at steady state each measured variables both 2 reactors. While the measured adjust in order to be more accurate than the measurement.

In this case, the measured mass flow rates in process have 6 variables: $F_{W,in}, F_{CL,in}, F_{W,v,1}\left(vapor\ phase\right), F_{pre-polymer,1}, F_{W+CL,v,2}\left(vapor\ phase\right), F_{nylon6,2}$ as show in Figure 4.18. Let us also ignore the energy flows of this process and focus only on the mass flow rates

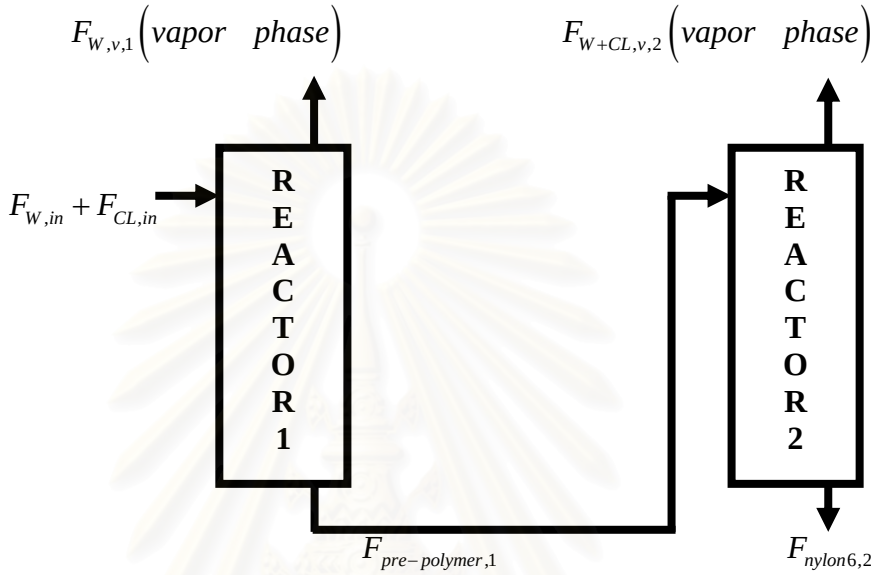


Figure 4.18 Mass flow rates diagram both 2 reactors for industrial nylon 6 production process

Therefore, the mass flow rates balance around the reactor 1 and reactor 2 can be written as:

Reactor 1:

$$F_{W,in}+F_{CL,in}-F_{W,v,1}-F_{pre-polymer,1}=0 \tag{4.20}$$

Reactor 2:

$$F_{pre-polymer,1}-F_{W+CL,v,2}-F_{nylon6,2}=0 \tag{4.21}$$

The measured all mass flow rate values both 2 reactors do not satisfy in equation 4.20 and 4.21. The problem in this case is to reconcile all the mass flows so as to satisfy material balance of reactor 1 and reactor 2. In addition, it is required to estimate true value of each measured mass flow rates. Therefore, we take data reconciliation with gross error detection apply this process and find algorithm that

suitable for industrial nylon 6 production process. Reconciliation solution for the measured mass flow rates. (shown in Table 4.3)

Table 4.3 Reconciliation solution for the measured variables in industrial nylon 6 production process at steady state condition (both 2 reactors)

Mass flow rate of each component measured (kg/hr)	Algorithms for data reconciliation			
	WLS	CN	Lorentzian	hampel
water inlet	1384.990832	1384.990821	1384.990692	1384.991
caprolactam inlet	21.23286002	21.23291696	21.23343858	21.23286
water (vapor)	7.912654701	7.912671073	7.91281431	7.912655
Pre-polymer	1398.311037	1398.311067	1398.311316	1398.311
water+caprolactam(vapor)	13.01850506	13.01851729	13.01861824	13.01851
polymer	1385.292532	1385.292549	1385.292698	1385.293

As a result of reconciled data each algorithm, the case apply data reconciliation with gross error detection in industrial nylon 6 production process both 2 reactors. It can be seen that reconciled data obtained from 4 algorithms are close. It can also be said that measurement data both 2 reactor have Normal distribution (only random error). Due to the performance test in case A and case B, the reconciled data obtained from Weighted Least-Square method is close to true value at steady state condition if the measurement data have only random error. Therefore, algorithms are the appropriate in this case: Weighted Least-Square, Contaminated Normal distribution function, Lorentzian distribution function and Hampel's redescending M-estimator approaches. Because, the reconciled data obtained from 4 algorithms are close to true value at steady state condition. However in case measurement data have gross errors present in measurement, Weighted Least-Square method can not support with gross errors present in measurements. Therefore, Robust function method can perform better than Weighted Least-Square method because it can support with process measurements contain both gross and random errors. In addition, the performance of Robust function method: Contaminated Normal, Lorentzian distribution function and Hampel's redescending M-estimator approaches are same depend on the tuning parameter of each algorithm.

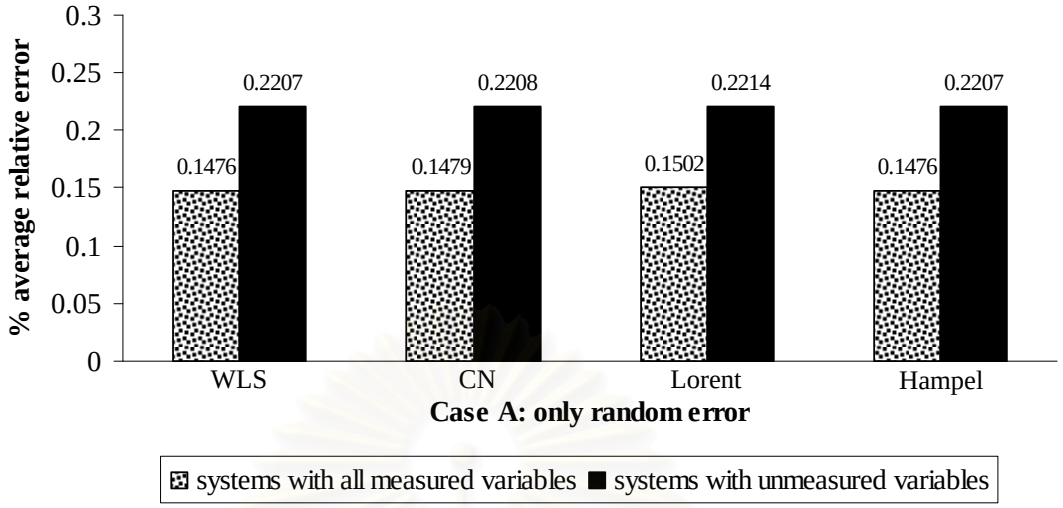
4.2.2 Systems with Unmeasured Variables

In the previous example, we have assumed that all variables are measured. However, usually only a subset of the variables is measured. The presence of unmeasured variables not only complicates the problem solution, but also introduces new questions such as whether an unmeasured variable can be estimated, or whether a measured variable can be reconciled as illustrated by the following example.

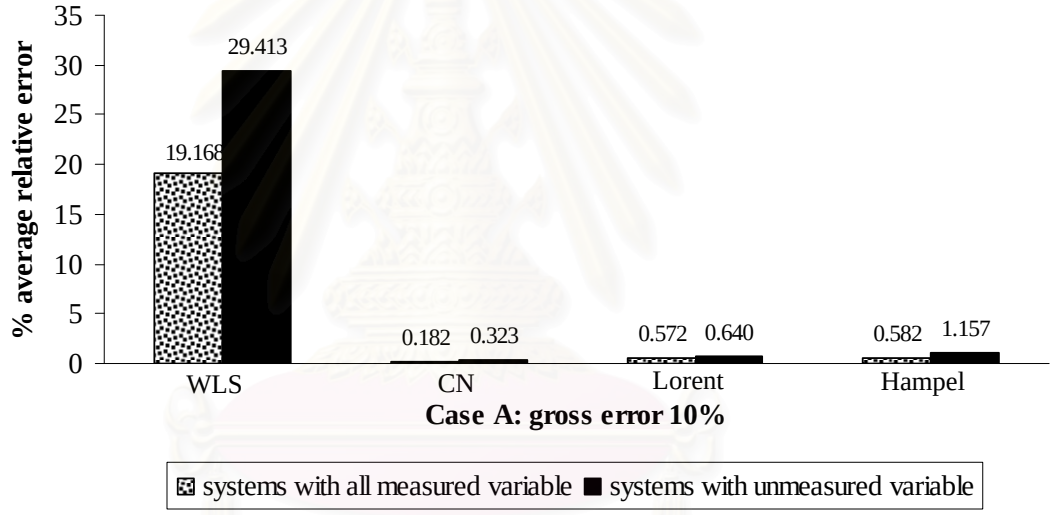
Let us consider the flow reconciliation problem of the case A, case B and case C are studied in the previous, we are assume:

- Case A: flows of variables A, B are measured, while the C is unmeasured.
- Case B: flows of in/outlet water, in/outlet caprolactam and cyclic dimer are measured, while the nylon 6 is unmeasured.
- Case C: flows of all variable in process are measured, while the $F_{W,v,1}\left(vapor \quad phase\right)$ and $F_{W+CL,v,2}\left(vapor \quad phase\right)$ are unmeasured.

Thereafter, we taking data reconciliation with gross error detection each algorithm applies to system with unmeasured variables in case A, case B and case C respectively. To estimate current status of process variables and unmeasured variables in all case. The efficiency of system, as applied data reconciliation for estimate unmeasured variables are shown %average relative error. The comparison between system with unmeasured variables and system with all measured variables are shown in Figure 4.19 - Figure 4.21.

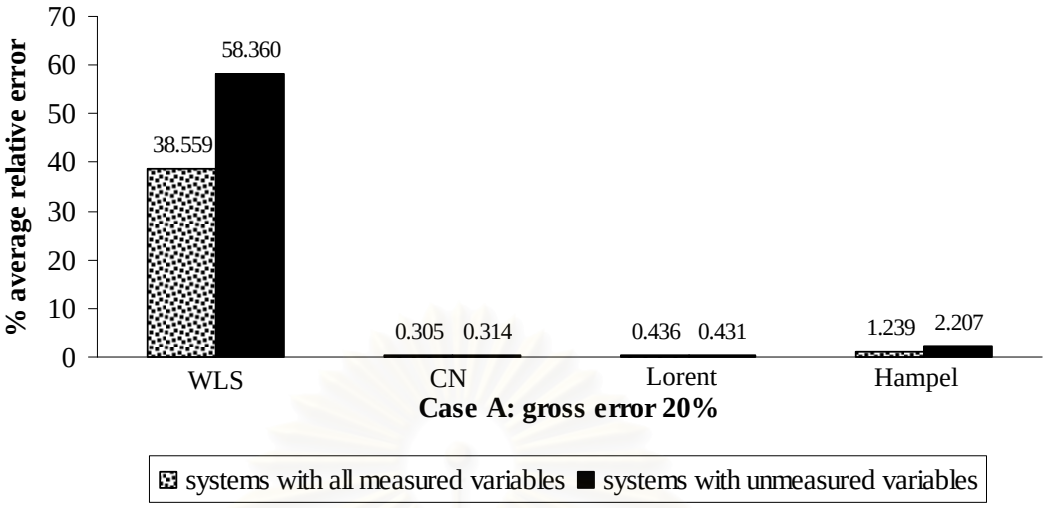


(a)

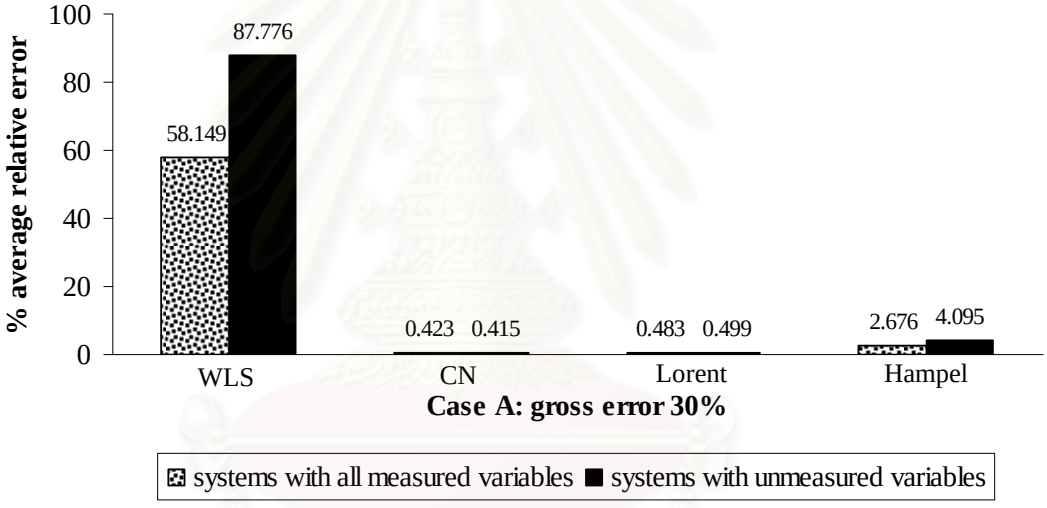


(b)

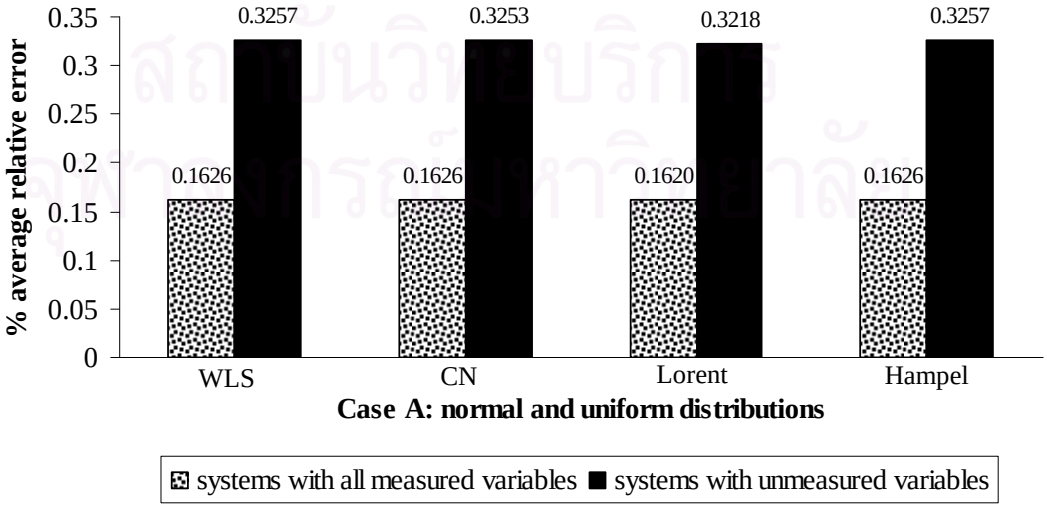
Figure 4.19 %Average relative error in all cases of applying data reconciliation with gross error detection in case A: measurement data have only random error (a), measurement data have gross error 10% (b), measurement data have gross error 20% (c), measurement data have gross error 30% (d) and measurement data contain both normal and uniform distributions (e).



(c)

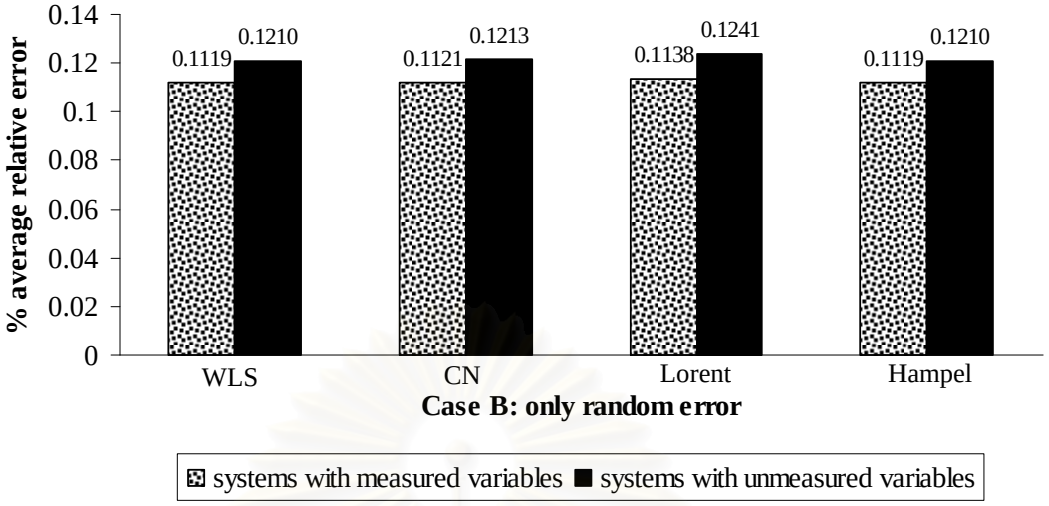


(d)

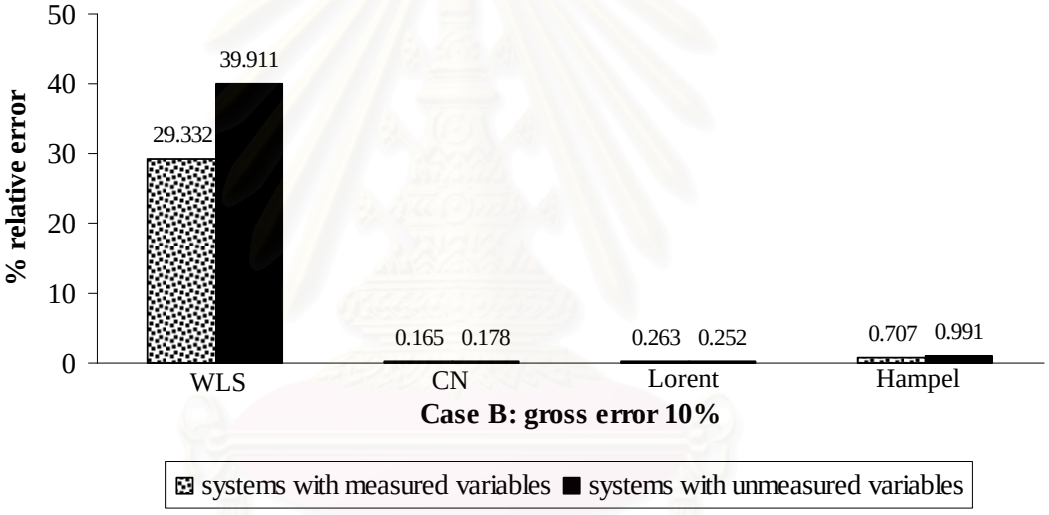


(e)

Figure 4.19 (continued)

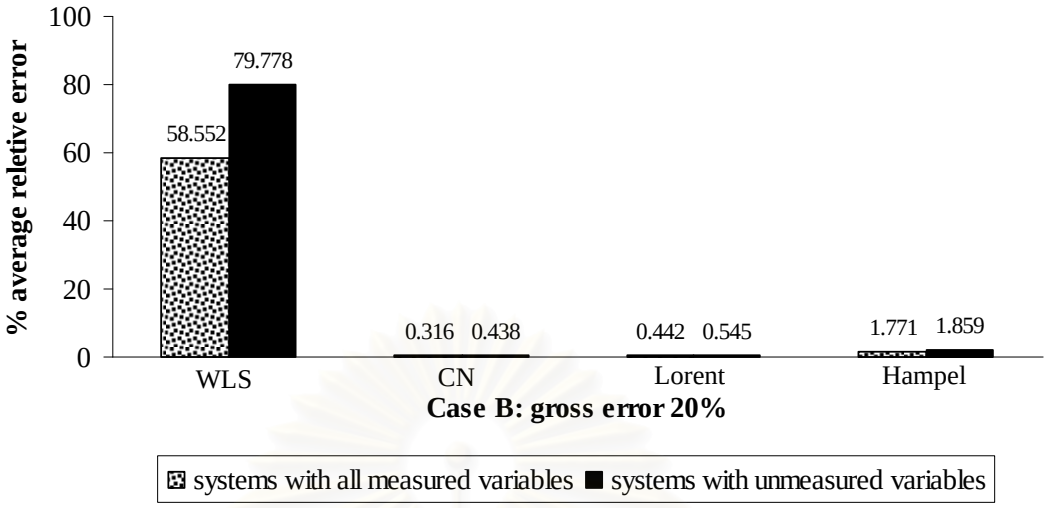


(a)

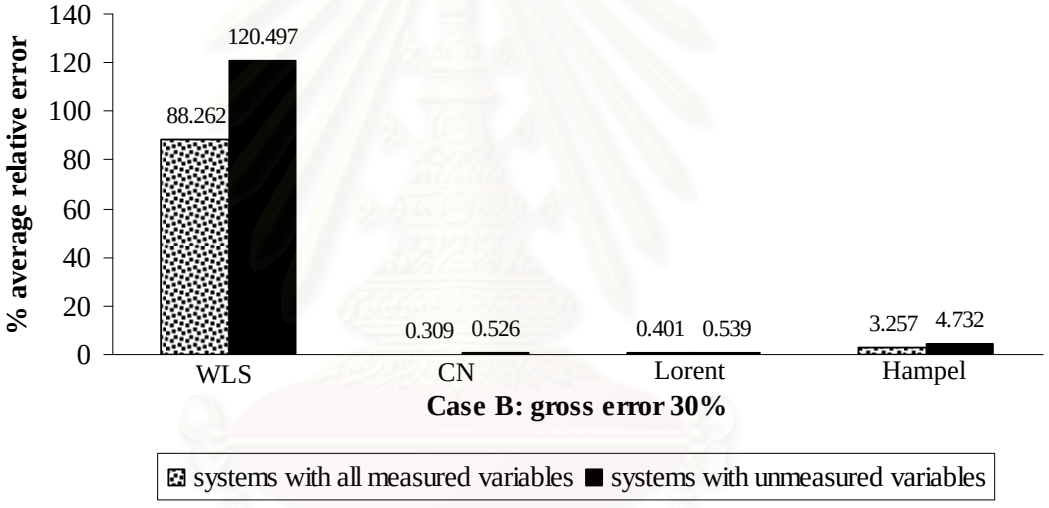


(b)

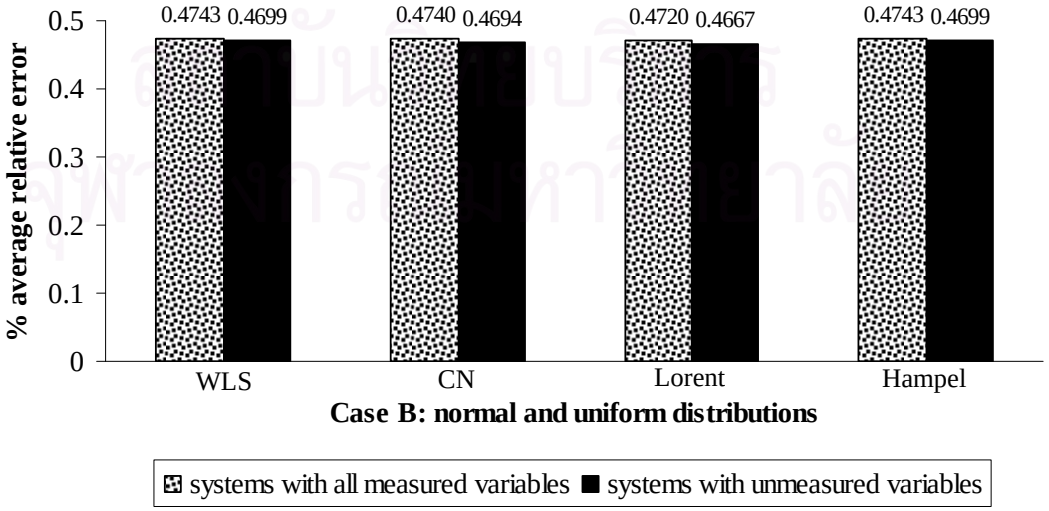
Figure 4.20 %Average relative error in all cases of applying data reconciliation with gross error detection in case B: measurement data have only random error (a), measurement data have gross error 10% (b), measurement data have gross error 20% (c), measurement data have gross error 30% (d) and measurement data contain both normal and uniform distributions (e).



(c)



(d)



(e)

Figure 4.20 (continued)

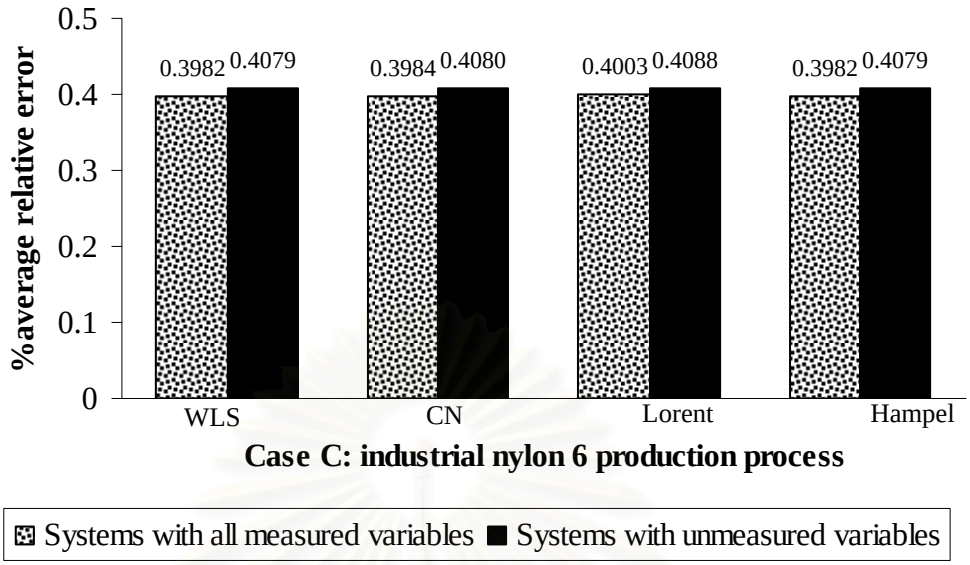


Figure 4.21 %Average relative error of applying data reconciliation with gross error detection in case C

In Figure 4.19 - Figure 4.21, demonstrate the comparison of %average relative error of solution obtained from data reconciliation in case A, B and C both case of systems with all measured variables and systems with unmeasured variables. As a result in Figure 4.19 and Figure 4.20, it can be seen that %average relative error have stronger tendency to be increase with an increase amount of gross error. In addition, the Weighted Least-Square method has highest %average relative error if gross errors presented in measurement data. On the other hand, Contaminated Normal distribution function, Lorentzian distribution function and Hampel's redescending M-estimator approaches show a better performance. In case measurement data have only random error. In the same way, %average relative error obtained from 4 algorithms can estimate unmeasured variables. Therefore, algorithms are the appropriate in case A and case B: Contaminated Normal distribution function, Lorentzian distribution function and Hampel's redescending M-estimator approaches. Because it can support with process measurements contain both gross and random errors.

As shown in Figure 4.21, it can be seen that %average relative error all algorithms of system with unmeasured variables close to system with all measured variables. It is also common to say that unmeasured variables can be estimated by data reconciliation, providing that enough measured data is available in order to make

them observable. These algorithms also work well in the case where unmeasurement process streams exist.

CHAPTER V

CONCLUTIONS AND RECOMMENDATIONS

5.1 Conclusion

Data reconciliation is an important step in real time on-line optimization. It adjusts the process measurements with random errors to satisfy the constraints of the system model and provides estimates for unmeasured variables and process parameters, which are used in the consecutive economic optimization step for process near its optimum condition.

In this research, we implement simultaneous data reconciliation and gross error detection strategies to industrial nylon 6 production process at steady state condition for cross-check of real time process measurements. We study the performance of each available technique that suitable for industrial nylon 6 production process and compare these algorithms under the conditions where some process streams are unmeasured. The algorithms that we studied are as follows: Weighted Least-Square, Contaminated Normal**,** Lorentzian distribution function and Hampel's redescending M-estimator.

Before, we study the performance of each available technique that suitable for industrial nylon 6 production process. We taking data reconciliation with gross error detection apply to other cases for compare performance each algorithm. There are three cases: Application data reconciliation with gross error detection in simple case, nylon 6 production process by simulated at steady state condition and industrial nylon 6 production process.

Result of application data reconciliation with gross error detection in all cases. The result shows that in the case of process measurements contain only random error. The reconciled data obtained from 4 algorithms are close to true value because measurement data of each variable in process have normal distribution. While most of

measurement values are distributed over mean range of each variables measured. Therefore, all algorithms perform well in this case.

For comparison, in the case that process measurements contain both gross and random errors, it can be seen that robust function method: Contaminated Normal distribution function, Lorentzian distribution function and Hampel's redescending M-estimator show a better performance. The reconciled data obtained each algorithms are close to true value. Each algorithm can support with gross errors present in measurements depending on adjusted tuning parameter of each algorithms. Therefore, with 10%, 20% or 30% of gross errors in measurement does not affect to reconciled data obtained. On the other hand, the reconciled data obtained from WLS approach are different to true value of each variables because can not detect gross error. We conclude that WLS is appropriate for measurement data containing only random error.

Therefore, apply data reconciliation with gross error detection in VK column reactors for industrial nylon 6 production process at steady state condition. Algorithms are the appropriate in this case: Contaminated Normal distribution function, Lorentzian distribution function and Hampel's redescending M-estimator approaches. Because it can support with process measurements contain both gross and random errors. Moreover, these algorithms also work well in the case where unmeasurement process streams exist.

5.2 Recommendations

The developed system (data reconciliation with gross error detection techniques) is implemented in distributed control system (DCS) with real time optimization in industrial nylon 6 production process.

REFERENCES

Chouaib, B., Data Reconciliation as a Framework for Chemical Processes Optimization and Control. Doctor of Science, Department of Chemical Engineering, Faculty of Engineering Universitat Politcnica de Catalunya, 2004.

Jain, S. L., and Sharma, Dr. N. D., Chemistry of Nylon-6 Part III: Hydrolytic Polymerization of Caprolactam. Sir Padampat Research Centre, J. K. Nagar, KOTA-324 003(Raj), MAN-MADE TEXTILES IN INDIA, 1997.

Jain, S. L., and Sharma, Dr. N. D., Chemistry of Nylon-6 Part IV :(Technical Aspects of Hydrolytic Polymerization of Caprolactam). Sir Padampat Research Centre, J. K. Nagar, KOTA-324 003 (Raj), MAN-MADE TEXTILES IN INDIA, 1997.

Narasimhan, S., and Jordache, C., Data reconciliation and gross error detection: An intelligent use of process data. Houston, TX: Gulf Publishing Company, New York, 2000.

Pornsiri, K., Application of data reconciliation for steel pickling process. Master of Engineering Thesis, Department of Chemical Engineering, Faculty of Engineering Chulalongkorn University, 2002.

Seavey, K. C., Research development of simulation and optimization technology for commercial nylon-6 manufacturing processes. Doctor of Philosophy, Department of Chemical Engineering, Faculty of the Virginia Polytechnic Institute and State University, 2003.

Arora, N., and Biegler, L. T., Redescending estimators for data reconciliation and parameter estimation. Computers and Chemical Engineering. 25 (2001):1585–1599.

Chen, X., Pike, R. W., Hertwig, T. A., and Hopper, J. R., Optimal implementation of on-line optimization. Computer and Chemical Engineering. 22 (1998): 435–442.

Crowe, C.M., Data reconciliation-progress and challenges. Journal of Process Control. 6 (1996): 89–98.

Crowe, C. M., Reconciliation of process flow rates by matrix projection. Part II: The nonlinear case. AIChE Journol. 32 (1986): 616-623.

Devanathan, S., Vardeman, S. B., and Rollins, D. K., Likelihood and Bayesian methods for accurate identification of measurement biases in pseudo steady-state processes. Chemical Engineering Research and Design. 83 (2005): 1391–1398.

Johnston, L. P. M., and Kramer, M. A., Maximum likelihood data rectification: Steady-state systems. AIChE J. 41 (1995): 2415–2426.

Kuehn, D. R., and Davidson, H., Computer control. II: mathematics of control. Chem. Eng. Prog. 57 (1961): 6.

Myung, I. J., Tutorial on maximum likelihood estimation. Journal of Mathematical Psychology. 47 (2003): 90–100.

Narasimhan, S., and Mah, R. S. H., Generalized Likelihood Ratios for gross error identification in dynamic processes. AIChE J. 34 (1988): 1321.

Özyurt, D.B., and Pike, R.W., Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. Comput. Chem. Eng. 28 (2004): 381–402.

Rollins, D. K., and Davis, J. F., Unbiased Estimation of Gross Errors in Process Measurements. AIChE J. 38 (1992): 563.

Rosenberg, J., Mah, R. S. H., and Jordache, C., Evaluation of Schemes for Detecting and Identification of Gross Errors in Process Data. Ind. & Eng. Chem. Proc. Des. Dev. 26 (1987): 555–564.

Seavey, K. C., Khare , N. P., and Lui, Y. A., A new phase equilibrium model for simulating industrial nylon-6 production trains. Ind. Eng. Chem. Res. 42 (2003): 3900-3913.

Storn, R., Differential evolution: A simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization. 11 (1997): 341-359.

Tjoa, I.B., and Biegler, L.T., Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems. Compute. Chem. Eng. 15 (1991): 679–690.

Zhang, Z., and Pike, R. W., Source reduction from chemical plants using on-line optimization. Waste Management. 15 (1995): 183-191.

APPENDICES

APPENDIX A

REACTIONS AND KINETICS FOR NYLON 6 POLYMERIZATION

When we say that polymeric molecules are made up of multiple functional groups, we mean that functional group segments are connected in a linear chain by covalent bonds.

There are two types of functional group segments: bound (or repeat) segments and terminal (or end group) segments (Figure A.1).

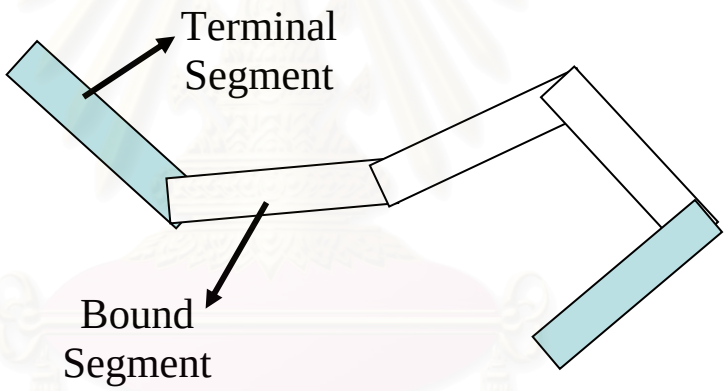


Figure A.1 A five segment, linear polymer chain consisting of two terminal segments and three bound segments (Kevin , 2003)

Terminal segments are found only at the ends of polymer chains, and are connected to other segments through one covalent bond. Bound segments, on the other hand, occur in the interior of a polymer molecule and have two covalent bonds.

Nylon-6 segments include the nylon-6 repeat segments (B-ACA) and the end groups terminal amine (T-NH2), terminal carboxylic acid (T-COOH) illustration as Figure A.2

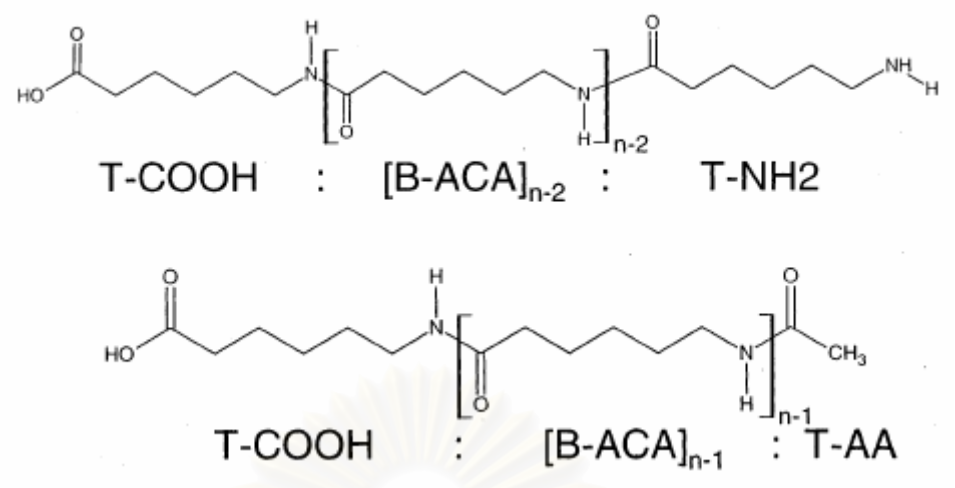


Figure A.2 Nylon-6 molecules of degree of polymerization *n* existing as two types: unterminated (above) and terminated by AA (below). (Kevin, Neeraj & Liu, 2003)

For Nylon-6 polymerization reactions and rate constants, we can follow as (Kevin, Neeraj & Liu, 2003) (presented in Table A.1 and Table A.2)

Table A.1 Nylon-6 polymerization reactions written in segment notation

equilibrium reaction	reaction rate
Ring opening of caprolactam $(W + CL \underset{k_1^{'}}{\overset{k_1}{\rightleftarrows}} P_n)$	
$CL + W \underset{k_1^{'}=k_1/K_1}{\overset{k_1}{\rightleftarrows}} P_1$	$R_1 = k_1[CL][W] - k_1^{'}[P_1]$
Polycondensation $(P_m + P_n \underset{k_2^{'}}{\overset{k_2}{\rightleftarrows}} P_{m+n} + W)$	
$P_1 + P_1 \underset{k_2^{'}=k_2/K_2}{\overset{k_2}{\rightleftarrows}}$ $T-COOH : T-NH_2 + W$	$R_2 = k_2[P_1]^2 - k_2^{'}[P_1][W]$
$P_1 + T-COOH \underset{k_2^{'}=k_2/K_2}{\overset{k_2}{\rightleftarrows}}$ $T-COOH : B-ACA + W$	$R_3 = k_2[P_1][T-COOH]$ $-k_2^{'}[W][T-COOH]\left(\frac{[B-ACA]}{[B-ACA]+[T-NH_2]}\right)$
$T-NH_2 + P_1 \underset{k_2^{'}=k_2/K_2}{\overset{k_2}{\rightleftarrows}}$ $T-NH_2 : B-ACA + W$	$R_4 = k_2[T-NH_2][P_1]$ $-k_2^{'}[W][T-NH_2]\left(\frac{[B-ACA]}{[B-ACA]+[T-COOH]}\right)$

$T-NH_2+T-COOH \underset{k_2^{'}=k_2/K_2}{\overset{k_2}{\rightleftarrows}}$ $B-ACA : B-ACA+W$	$R_5 = k_2[T-NH_2][T-COOH]$ $-k_2^{'}[W][B-ACA]\left(\frac{[B-ACA]}{[B-ACA]+[T-NH_2]}\right)$
equilibrium reaction	reaction rate
Polyaddition of caprolactam $(CL+P_n \underset{k_3^{'}}{\overset{k_3}{\rightleftarrows}} P_{n+1})$	
$P_1+CL \underset{k_3^{'}=k_3/K_3}{\overset{k_3}{\rightleftarrows}}$ $T-NH_2 : T-COOH$	$R_6 = k_3[P_1][CL]$ $-k_3^{'}[T-COOH]\left(\frac{[T-NH_2]}{[B-ACA]+[T-NH_2]}\right)$
$T-NH_2+CL \underset{k_3^{'}=k_3/K_3}{\overset{k_3}{\rightleftarrows}}$ $T-NH_2 : B-ACA$	$R_7 = k_3[T-NH_2][CL]$ $-k_3^{'}[T-NH_2]\left(\frac{[B-ACA]}{[B-ACA]+[T-COOH]}\right)$
Ring opening of cyclic dimer $(W+CD \underset{k_4^{'}}{\overset{k_4}{\rightleftarrows}} P_2)$	
$CD+W \underset{k_4^{'}=k_4/K_4}{\overset{k_4}{\rightleftarrows}}$ $T-COOH : T-NH_2$	$R_8 = k_4[CD][W]$ $-k_4^{'}[T-COOH]\left(\frac{[T-NH_2]}{[B-ACA]+[T-NH_2]}\right)$
Polyaddition of cyclic dimer $(CD+P_n \underset{k_5^{'}}{\overset{k_5}{\rightleftarrows}} P_{n+2})$	
$P_1+CD \underset{k_5^{'}=k_5/K_5}{\overset{k_5}{\rightleftarrows}}$ $T-NH_2 : B-ACA : T-COOH$	$R_9 = k_5[P_1][CD]$ $-k_5^{'}[T-COOH]\left(\frac{[B-ACA]}{[B-ACA]+[T-NH_2]}\right)$ $\left(\frac{[T-NH_2]}{[B-ACA]+[T-NH_2]}\right)$
$T-NH_2+CD \underset{k_5^{'}=k_5/K_5}{\overset{k_5}{\rightleftarrows}}$ $B-ACA : B-ACA : T-NH_2$	$R_{10} = k_5[T-NH_2][CD]$ $-k_5^{'}[T-NH_2]\left(\frac{[B-ACA]}{[B-ACA]+[T-COOH]}\right)$

Table A.2 Rate constants for the equilibrium reactions in Table A.1

rate constant expression	$k_i = A_i^0 \exp\left(-\frac{E_i^0}{RT}\right) + A_i^c \exp\left(-\frac{E_i^c}{RT}\right)[T-COOH]$
equilibrium constant expression	$K_i = \frac{k_i}{k_i^{'}} = \exp\left(\frac{\Delta S_i - \Delta H_i / T}{R}\right)$

i	A_i^0 (kg/mol*s)	E_i^0 (J/mol)	A_i^c (kg^2/mol^2*s)	E_i^c (J/mol)	ΔH_i (J/mol)	ΔS_i J/mol*K
1	1.66E+02	8.32E+04	1.20E+04	7.87E+04	8.03E+03	-33.01
2	5.26E+06	9.74E+04	3.37E+06	8.65E+04	-2.49E+04	3.951
3	7.93E+05	9.56E+04	4.55E+06	8.42E+04	-1.69E+04	-29.08
4	2.38E+08	1.76E+05	6.47E+08	1.57E+05	-4.02E+04	-60.79
5	7.14E+04	8.92E+04	8.36E+05	8.54E+04	-1.33E+04	2.439

We can define molecular weights of each functional group as Table A.3

Table A.3 Functional group molecular weights

Functional Group	Molecular Weight (g/mol)
W	18.01528
CL	113.1595
CD	226.318
ACA	131.1742
T-COOH	130.1668
T-NH2	114.1674
B-ACA	113.1595

APPENDIX B

RECONCILIATION SOLUTION FOR THE MEASURED VARIABLES IN ALL CASES

Table B.1 Reconciliation solution for the measured variables in simple case (Case 1: No gross errors present in measurements)

variables	True values	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
A	1.00	1.001018	1.001024	1.00107	1.001018
B	2.00	1.995498	1.995493	1.995448	1.995498
C	3.00	2.996517	2.996517	2.996518	2.996517

Table B.2 Reconciliation solution for the measured variables in simple case (Case 2: have gross errors present in measurements 10 %)

variables	True values	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
A	1.00	1.268121	1.000808	1.008867	1.0125
B	2.00	2.261061	1.994084	1.986489	1.9952
C	3.00	3.529182	2.994892	2.995356	3.0077

Table B.3 Reconciliation solution for the measured variables in simple case (Case 2: have gross errors present in measurements 20 %)

variables	True values	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
A	1.00	1.538987	1.001132	1.003791	1.0277
B	2.00	2.525761	1.989931	1.987331	1.9985
C	3.00	4.064748	2.991063	2.991122	3.0262

Table B.4 Reconciliation solution for the measured variables in simple case (Case 2: have gross errors present in measurements 30 %)

variables	True values	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
A	1.00	1.809981	1.005701	1.007499	1.0529
B	2.00	2.797385	1.98934	1.988597	2.0117
C	3.00	4.607366	2.995041	2.996097	3.0646

Table B.5 Reconciliation solution for the measured variables in simple case (Case 3: measurement data contain both Normal and Uniform distributions)

variables	True values	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
A	1.00	0.999872	0.999872	0.999864	0.999872
B	2.00	2.005753	2.00575	2.005723	2.005753
C	3.00	3.005625	3.005621	3.005587	3.005625

Table B.6 Reconciliation solution for the measured variables in nylon 6 production process by simulated at steady state condition (Case 1: No gross errors present in measurements)

Mass flow rate of each component measured (kg/hr)	True values at steady state condition	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
Water	1.00	1.004922	1.004929	1.004989	1.004922
Caprolactam	99.00	99.00585	99.00585	99.00586	99.00585
Water	0.845419	0.846364	0.84637	0.846417	0.846364
Caprolactam	23.3329	23.33496	23.33495	23.33489	23.33496
Cyclic dimer	0.316816	0.31695	0.31695	0.316946	0.31695
Nylon 6	75.50477	75.5125	75.51251	75.5126	75.5125

Table B.7 Reconciliation solution for the measured variables in nylon 6 production process by simulated at steady state condition (Case 2: have gross errors present in measurements 10%)

Mass flow rate of each component measured (kg/hr)	True values at steady state condition	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
Water	1.00	1.54095	1.006828	1.013279	1.0257
Caprolactam	99.00	99.54105	99.00245	99.00188	99.0116
Water	0.845419	1.112937	0.847639	0.846681	0.8459
Caprolactam	23.3329	23.58185	23.32843	23.3312	23.3393
Cyclic dimer	0.316816	0.596441	0.316849	0.317033	0.3217
Nylon 6	75.50477	75.79077	75.51636	75.52025	75.5304

Table B.8 Reconciliation solution for the measured variables in nylon 6 production process by simulated at steady state condition (Case 2: have gross errors present in measurements 20%)

Mass flow rate of each component measured (kg/hr)	True values at steady state condition	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
Water	1.00	2.07879	1.004693	1.00722	1.0508
Caprolactam	99.00	100.0823	99.00615	99.00569	99.0339
Water	0.845419	1.38279	0.848227	0.850463	0.8594
Caprolactam	23.3329	23.83069	23.32681	23.3278	23.3523
Cyclic dimer	0.316816	0.874071	0.313529	0.312747	0.3286
Nylon 6	75.50477	76.07359	75.52228	75.5219	75.5444

Table B.9 Reconciliation solution for the measured variables in nylon 6 production process by simulated at steady state condition (Case 2: have gross errors present in measurements 30%)

Mass flow rate of each component measured (kg/hr)	True values at steady state condition	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
Water	1.00	2.630183	1.006127	1.008023	1.0726
Caprolactam	99.00	100.6183	99.00575	99.00538	99.0723
Water	0.845419	1.651885	0.846223	0.848195	0.8706
Caprolactam	23.3329	24.08994	23.33341	23.33386	23.3733
Cyclic dimer	0.316816	1.156899	0.313266	0.312862	0.3453
Nylon 6	75.50477	76.34979	75.51898	75.51848	75.5557

Table B.10 Reconciliation solution for the measured variables in nylon 6 production process by simulated at steady state condition (Case 3: measurement data contain both Normal and uniform distributions)

Mass flow rate of each component measured (kg/hr)	True values at steady state condition	Algorithms for data reconciliation			
		WLS	CN	Lorentzian	Hampel
Water	1.00	0.999788	0.999787	0.99978	0.999788
Caprolactam	99.00	99.00671	99.00671	99.00669	99.00671
Water	0.845419	0.84605	0.846056	0.846104	0.84605
Caprolactam	23.3329	23.32773	23.32774	23.32787	23.32773
Cyclic dimer	0.316816	0.325436	0.32543	0.325373	0.325436
Nylon 6	75.50477	75.50728	75.50726	75.50713	75.50728

VITA

Mr. Jakapon Veeravong was born August 28, 1983 in Kanchanaburi, Thailand. He graduated Bachelor Degree of Engineering in Chemical Engineering from Srinakharinwirot University in 2005. After that he studied for master degree in Chemical Engineering and joined Control and Systems Engineering research group at Chulalongkorn University in 2006.