

Mälardalen University Licentiate Thesis

No.34

TOOLS FOR RECONCILIATION OF MEASUREMENT DATA FOR PROCESSES AT STEADY-STATE

Christer Karlsson

2004



Department of Public Technology
Mälardalen University

Copyright © Christer Karlsson, 2004

ISSN number: 1651-9256

ISBN number: 91-88834-60-3

Printed by Arkitektkopia, Västerås, Sweden

Distribution: Mälardalen University Press

Abstract

Reliable sensor readings are important for running a heat and power plant. Process and sensors degenerate during plant operation. This can be handled by on-line data treatment. It works in three steps; first large errors in measurements are detected, then the errors are isolated and removed. The third step is general data estimation that both reconcile data by using redundant data and estimates sensor values.

In the first part of the work (report 1) in this thesis, the hypothesis that data estimation can be done despite degradation in both process and sensors was tested. The focus was on systematic errors in sensors. Three different models were theoretically evaluated for data treatment purposes. A first principles model was chosen and the data treatment method was built with the components; classification, detection, isolation and data reconciliation.

Extensive testing with a flue gas channel process model resulted in a second work (paper 1). The testing revealed drawbacks considering estimation of the size of the gross error and problems handling different magnitudes in connected mass flows.

The third work (paper 2) was an article on Bayesian networks (BN) for decision support on soot blowing superheaters in a bio-fuelled boiler. The aim was to construct a tool for prediction of situations where abnormal fouling is at risk and to give advice on preventive actions. To test if qualitative and quantitative methods could benefit from each other, a combination of first principles models and Bayesian networks was built. It turned out that the combination was necessary. The problem would have been sub-optimal if only BN had been used for modelling and the alternative with only first principles models would not have been possible with the resources at hand.

Fault isolation is one of the key components in a data treatment where gross errors are present. Analytically solving the estimation of a gross error had shown being difficult (report 1 and paper 1). A novel approach for isolation was therefore proposed in the fourth part of the work (paper 3). An optimisation approach with time window and penalty multiplier was tested for isolation. It turned out to work excellently.

The report and the three papers are complemented with a literature review. Each main method is followed by a selected method showing basic computations and principles of data treatment.

Preface

This thesis has been carried out at the Department of Public Technology, Mälardalen University, in collaboration with; Thermal Engineering Research Institute (Värmeforsk), The Swedish Energy Agency, ABB and Fortum.

I would like to express my gratitude to:

My supervisor Erik Dahlquist for inspiration, guidance, wild ideas and always having energy and a smile on his face to share. Erik Dotzauer for structure, realism and many hours of discussions on optimisation, reviewing and ideas in articles. Kimmo Eriksson for helping me with details in linear algebra. Carl-Fredrik Lindberg for the insights in a new method for gross error isolation. Finn Jensen for expert advice on Bayesian networks. The colleagues at the Department of Public Technology who have helped and been a support on both the technical and the social level. To all near and dear friends and family for always supporting and encouraging me to continue this adventure.

List of Appended Papers

This thesis is based on the following three papers and technical report:

Papers

- Karlsson C., Dahlquist E. and Dotzauer E., “Data Reconciliation and Gross Error Detection for the Flue Gas Train in a Heat and Power Plant”. *Published in conference on Probabilistic Methods Applied to Power Systems, USA, 2004.*
- Widarsson B., Karlsson C. and Dahlquist E., “Bayesian Network for Decision Support on Soot Blowing Superheaters in a Biomass Fuelled Boiler”. *Published in conference on Probabilistic Methods Applied to Power Systems, USA, 2004.*
- Karlsson C. and Dotzauer E., “Gross Error Isolation by Optimisation with Penalty Function”. *Submitted for publication in Computers and Chemical Engineering, Elsevier.*

Report

- Karlsson C. and Dahlquist E., “Process and sensor diagnostics – Data reconciliation for a flue gas channel”. *Värmeforsk Project Publication Series. 2003, P9-816.*

Nomenclature

A	Linear flow process model
A_1	Set of equations with measured variables
A_2	Set of equations with unmeasured variables
ANN	Artificial neural networks
AVTI	Average number of Type I error
BN	Bayesian networks
C	Incidence matrix
c	An element in incidence matrix C
cov	Covariance
e	Normal distributed noise, added to true signal
g	Linearly independent columns of A_2 (rank of A_2)
GE	Gross error
GLR	General likelihood ratio
GT	Global test
I	Total number of samples
i	Counter of samples
L	Number of variables
LCT	Linear combination technique
M	Total number of rows
m	Counter of row
Σ	Measurement error covariance matrix
MILP	Mixed integer linear programming
MINLP	Mixed integer nonlinear programming
MT	Measurement test
N	Total number of columns
n	Counter of columns
NLP	Nonlinear programming
NT	Nodal test
OP	Overall power
PCA	Principal component analysis
pdf	Probability density function
P	Permutation matrix
P_u	Permutation matrix of undetermined mass flow
Q	Test quantity
Q_u	Part of Q connected to undetermined mass flow
r	Residual

$R_{i,u}$	Vector showing linear dependence between unmeasured determinable and unmeasured indeterminable variables
R_u	Part of R connected to undetermined mass flow
S	Data set of sensors containing gross errors
s_y	Standard deviation (letter index)
s_i	Measured signals (number index)
T	Total number of time samples
t	Counter of time samples
U	Data set
u	Unmeasured
u_d	Determinable unmeasured variables
u_i	Indeterminable unmeasured variables
V	Covariance matrix
v	Degrees of freedom
W	Weight matrix
w	Element in weight matrix W
x	Estimated state
x_n	Non-redundant variables in measured matrix A_1
x_r	Redundant variables in measured matrix A_1
y	Measured state
y_{true}	Flow vector without noise and gross error
y_{meas}	Measured flow vector
$y_{meas,GE}$	Measured flow vector containing gross error
$Z_{1-\beta/2}$	Modified critical value for multiple hypothesis tests
$Z_{critical}$	Theoretic critical value for chi-square distribution that do not contain gross errors
α	Dampening
α_{sig}	Significance level (limit of Type I error)
β	Modified significance level α_{sig} , when using multiple hypothesis tests
γ	Test quantity for chi-square-tests
Γ	Gamma probability density function
δ	Gross error
ε	Bounds on the process model residual
μ	Penalty function multiplier
$\bar{\mu}$	Mean value
μ_x	Mean value of x
μ_y	Mean value of y

Table of Contents

Abstract	I
Preface	II
List of Appended Papers.....	III
Nomenclature	IV
Table of Contents	VI
List of Figures	VIII
List of Tables.....	VIII
1 INTRODUCTION.....	1
1.1. MOTIVES.....	3
1.2. RESEARCH APPROACH	4
1.3. PROBLEM DEFINITION	5
1.3.1. Cause of degradation in heat and power plant process and sensors	5
1.3.2. Effect of degradation in heat and power plant process and sensors	6
1.4. LITERATURE REVIEW	7
1.4.1. Detection	8
1.4.2. Isolation	9
1.4.3. Data reconciliation	10
2 DATA TREATMENT.....	12
2.1. PROCESS MODELS	12
2.2. SENSORS	14
2.3. CLASSIFICATION OF SENSORS	15
2.3.1. Example of classification by graph theory	17
2.3.2. Example of classification by matrix projection.....	17
2.4. SOME BASIC PROPERTIES USED IN STATISTICS	21
2.5. NORMAL AND CHI-SQUARE DISTRIBUTIONS	22
2.6. DETECTION OF GROSS ERRORS	25
2.6.1. Tests for detection	26
Global Test (GT)	26
Nodal Test (NT)	27
Measurement Test (MT).....	28
General Likelihood Ratio (GLR).....	28
2.6.2. Example of gross error detection by global test.....	28
2.7. ISOLATION	30
2.7.1. Isolation of single gross error based on statistical hypothesis test.....	31
2.7.2. Isolation of multiple gross errors based on statistical hypothesis test.....	32
2.7.3. Isolation of gross error by solving a sequence of nonlinear programming problems.....	33

2.7.4.	<i>Example of isolation by solving a sequence of NLP problems</i>	<i>34</i>
2.8.	DATA RECONCILIATION.....	35
2.8.1.	<i>Example of data reconciliation</i>	<i>37</i>
2.9.	APPLICATIONS USING TREATED DATA.....	39
3	RESULTS.....	40
3.1.	PAPER 1	40
3.2.	PAPER 2	43
3.3.	PAPER 3	45
3.4.	REPORT 1	47
4	CONCLUDING REMARKS.....	48
4.1.	CONCLUSIONS.....	48
4.2.	FUTURE WORK	49
5	REFERENCES	51

List of Figures

FIGURE 1. DATA TREATMENT STRUCTURE.	2
FIGURE 2. FUNDAMENTAL FACTORS IN SYSTEM RELIABILITY, [WAT89].	3
FIGURE 3. SIGNAL GRAPH. BLACK COLOURED SIGNALS ARE MEASURED. WHITE COLOURED SIGNALS ARE UNMEASURED.	17
FIGURE 4. FLOWSHEET DIAGRAM.	18
FIGURE 5. PROBABILITY DENSITY FUNCTION CURVES FOR NORMAL DISTRIBUTION. SOLID CURVE HAS $M=0$ AND DASHED CURVE HAS $M=2$	23
FIGURE 6. A CURVE FROM THE CHI-SQUARE PDF.	24
FIGURE 7. ISOLATION BASED ON STATISTICAL TESTS FOR SINGLE GROSS ERROR.	31
FIGURE 8. ISOLATION BASED ON STATISTICAL TESTS FOR MULTIPLE GROSS ERRORS.	32
FIGURE 9. ISOLATION BASED ON SOLVING A SEQUENCE OF NONLINEAR PROGRAMMING PROBLEMS.	33
FIGURE 10. DATA TREATMENT FROM RAW DATA TO APPLICATION.	41
FIGURE 11. BAYESIAN NETWORK FOR DECISION SUPPORT ON SOOT BLOWING.	44
FIGURE 12. EFFECT OF GROSS ERROR AND DAMPENING FOR A SINE WAVE. SOLID LINE IS TRUE SIGNAL, UPPER DASHED LINE IS SIGNAL WITH GROSS ERROR AND LOWER DASHED LINE IS DAMPENED SIGNAL.	45

List of Tables

TABLE 1. SENSOR READINGS WITH GROSS ERROR IN SENSOR 5 EQUAL TO -3 KG/S IN EACH SAMPLE.	34
TABLE 2. RESULT FROM SOLVING A SEQUENCE OF NLP PROBLEM FOR EACH SENSOR TO ISOLATE THE GROSS ERROR.	35
TABLE 3. SENSOR READINGS WITH CORRECTED GROSS ERROR IN SENSOR 5 EQUAL TO -2.97 KG/S.	38
TABLE 4. TWO-NORM OF THE RESIDUAL IN DIFFERENT STEPS OF THE DATA TREATMENT.	39
TABLE 5. SUMMARY OF PERFORMANCE MEASURES.	42

1 Introduction

Measuring properties are essential for plant control and monitoring. Constant strife for improved plant efficiency demand better control, which often means more instrumentation. The sensors are used for on-line computation of e.g. plant operation cost, emissions, heat power, etc. These applications need reliable on-line measurement data. A known problem is that sensors in the same location show different values. Regarding one as faulty or not reliable often solves this or an average computation is made. The use of simulators make it possible to both measure and compute the same state variable, it makes the situation complex when it comes to selecting reliable sensors and what sensor readings to believe in.

These problems can be solved by a method able to handle redundant information and estimate the state of the process. The three components; gross error detection, isolation and data reconciliation can be put together to form a system meeting the demands mentioned. The research area of data reconciliation and gross error detection spans over many research areas, such as signal analysis, statistics, optimisation, process control, information theory, etc. Complex industries have much to gain from data treatment. The development in the field is fastest in the nuclear power generation and the chemical process industry. An accident in these industries may be disastrous with consequence in both causalities and company bankruptcy, [Hoo04]. Other areas using this kind of data treatment are the mining industry and the conventional power generation industry.

The chain from sensor signal to operator interaction must be considered to find optimal solutions. The economic considerations on number of sensors and maintenance costs must also be included when choosing if, and by what degree the technique is to be implemented.

The core of data treatment in this thesis is data reconciliation (DR). DR algorithms can handle random noise in the measurements, but they cannot handle systematic and large errors. A protective layer of gross error detection and gross error isolation are therefore vital parts of a data treatment system. The goal of the data treatment is to deliver a data set, which is closer to the true process state than raw data. This data set is ready for use in applications on the next level in the hierarchy. Tools and techniques used for data treatment aiming to deliver state estimation using sensors measurements and process models under degradation is the topic of this thesis.

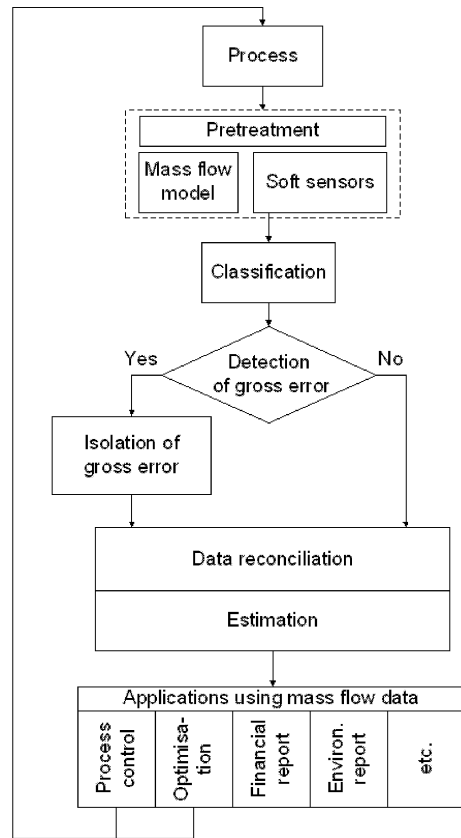


Figure 1. Data treatment structure.

In the thesis the common steps in data treatment are followed, see Figure 1, and each method is explained, but first the motives and the research approach are defined. A literature review on methods used until today is presented. Data treatment is broken down into pieces and explained part by part. Classification, detection, isolation and data reconciliation sections are closed by examples, to help the reader to follow the thread throughout the thesis. The same process model is used in all examples. Results are presented in form of summaries of the attribution of each article to the research task. Everything is then concluded and the future research outline is drawn.

For the reader not acquainted with the terms used in gross error detection and data reconciliation, but who is familiar with the Swedish language, a short wordlist is available in Report 1, in the appendix.

1.1. Motives

Plant safety is important for all process industry. It is a total concept to be implemented on every level. Some tasks in plant safety may be either fully automated or partially automated complemented with dialogue with plant operators. The main motive is to handle degradation in both sensors and process on-line by exploiting both sensor network and sensor signals together with a process model. Advantages are increased product quality and plant safety, less unplanned plant shutdowns and more knowledge about the process when implementing such a system. Chemical and nuclear plants are a potential threat for people and environment, thus data treatment is one of the steps to prevent accidents by monitoring plant measurement equipment and the overall process and sensor degradation.

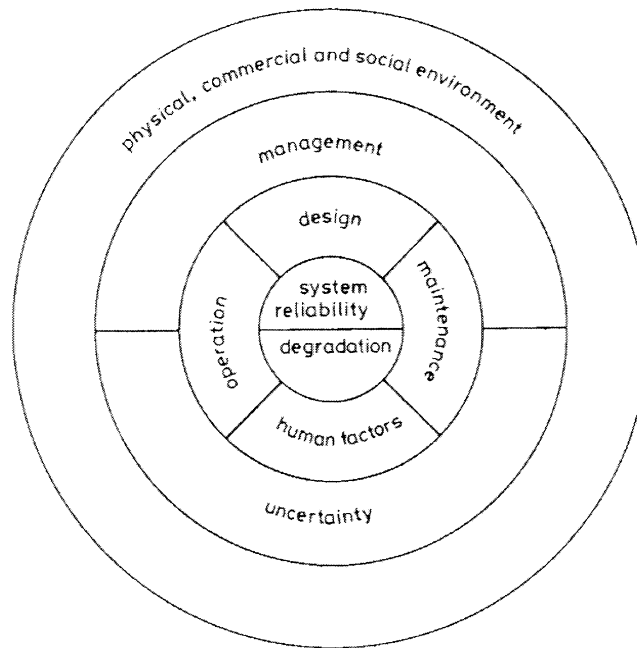


Figure 2. Fundamental factors in system reliability, [Wat89].

When components in Figure 2 fail then the effects are ranging from a not properly working plant to total disaster. Examples of accidents are found in e.g. the reports by Doyle [Doy72] and Kletz [Kle98] on recovery boilers. The Bhopal gas accident with more than 3000 casualties was the wake up call for implementing plant safety, [Hoo04]. In the context of plant safety, data treatment is the part that handles degradation in sensors and processes.

Diagnosis of a sensor as an individual is well developed in signal analysis, but there are much more information to be gained from the sensors. For example DR and gross error isolation techniques explores the qualitative information given in a sensor network and a process model together with the measured signals. The qualitative process model together with different levels of quantitative information given by e.g. standard deviations has potential as powerful tools for diagnosis.

In thermal processes, like heat and power plants, there are typically hundreds up to thousands of sensors installed for control, monitoring and alarm purposes. Some of these are maintained routinely. Why not all are in the maintenance program is due to the maintenance cost and also of lack of insight in what cost a faulty sensor can cause if it is used for control or as input for economical calculations.

No calibration and faulty installation is a cause for economical losses for process plants. Example of this is found in the paper by Wang et al., [Wan01], where data reconciliation is applied on a Chinese refinery resulting in reduced losses. It is the attractive concept of using (existing) process models together with existing sensors to increase plant availability and plant measurement usability that is exploited in this thesis to handle degradation in process and sensors.

1.2. Research approach

Hypothesis of this thesis:

Quantitative and qualitative methods can deliver a data set in balance from state estimation point a view, despite degradation in process and sensors.

To test this hypothesis the following tasks have been performed: a study was performed for three different process-modelling techniques and gross error detection techniques. Measurement data was gathered for major parts of the complete flue gas channel in a heat and power plant. Gross error detection and data reconciliation system was programmed and tested to be evaluated in depth. The system consists of first principles model, hypothesis test for gross error and isolation by sequential test. The complete data treatment system was programmed and evaluated. Positive and negative qualities were summarised in Report 1.

Sensitivity analysis on quality measures was performed on a modified heat and power plant scheme from Report 1. The results are presented in Paper 1. A combined approach of physical models and a Bayesian network was applied for the superheaters in a heat and power plant. The purpose was to investigate whether it is possible to predict how frequent soot-blowing of

the superheaters need to be performed to avoid a plant shutdown before annual revision. A secondary objective was to examine which measurements needed to be put into the Bayesian network and which needed to be put into physical models. A decision support for when to soot-blow was constructed and presented in Paper 2.

Two different formulations were tested for isolation of both gross error and dampening through isolation. The intention was to improve the performance of isolation and also to estimate both bias and dampening in the measured signal with a novel method proposed in Paper 3.

1.3. Problem definition

Measuring the true state of the process is essential for control and optimisation of a plant. In many processes the instrumentation and the process itself are affected by degradation and thus the measurements contain errors. Process models are difficult to maintain due to changes in process parameters combined with degeneration of the sensors. The problem of state estimation becomes interesting for any plant owner in the light of the costs for not running a plant optimally. Increasing operation safety by determining the plant state is also interesting, especially in nuclear power plants, but this is not in the scope of this thesis. Perhaps the greatest benefit is the possibility to decrease the number of unplanned plant shut-downs due to the tools provided. There are plants where degradation is not seen as one of the major problems, despite that these plants benefit from the possibilities of state estimation, classification of sensors and data reconciliation to reduce measurement noise.

The overall problem of handling degradation is interesting for the research community because of its complexity and challenges in integrating diverse methods like statistics, optimisation and modelling of thermal processes. Heat and power plants got all the challenging aspects and is therefore an interesting playground for trying new methods to solve the problem of handling degradation in sensors and the process. Thus it is possible to maintain a good estimation of the process state. Many other research areas such as advanced control, plant optimisation and functional maintenance greatly benefit from methods capable of estimating the true state of the process and estimating sensor readings of degrading or malfunctioning sensors.

1.3.1. Cause of degradation in heat and power plant process and sensors

Energy and environmental taxes and fees have pushed power plant owners to experiment with new fuels to lower costs. Fuel flexibility is more important than ever before, and it is

possible to meet this requirement with improved boiler technologies such as fluidised beds and circulating beds. Fuels earlier used only for heat water production is now introduced in heat and power plants, and these plants operate at higher temperatures. Also new chemical reactions are activated. The experience of combusting these new fuels and mixes of fuels in heat and power plants are increasing. Normal degradation of sensors and process in heat and power plants are mainly due to:

- Corrosion
- Fouling
- Erosion
- High temperature

Unbalance in plant heat balance or flue gas flow pattern cause problems. This can be due to use of fuels not intended or flaws in the plant design. The degradation causes listed above can force a shut down of the plant for maintenance. However, normally a certain degree of degradation is accounted for. Faster degradation in sensors is possible if they are positioned where they are exposed to erosion, corrosion or fouling. The same reasoning can be applied on the process if fuel quality differs from the expected, regarding for example moisture, alkali, sulphur, metals and properties such as; size distribution and burn-out rate.

Temperature sensors have a drift due to ageing and the high temperature they are working in. Positioning is important for temperature sensors in sections where large temperature gradients may occur. Pressure sensors are sensitive to clogging and also need careful positioning. Most plants are shut down and revised annually. During periodic maintenance and revision the process is controlled and the sensors are calibrated. Not all sensors are calibrated annually.

1.3.2. Effect of degradation in heat and power plant process and sensors

It is essential to control or reduce degradation of important sensors for alarm, control, optimisation and maintenance, primarily to reduce risk for personal injury, and secondly to minimise costs. Examples of effects from degradation are:

- Nuisance alarms, delayed alarms and not triggered alarms.
- Non-optimal controller set point and loss of control.
- Non-optimal overall plant optimisation.
- Faulty economic and environmental reporting.
- Damages on process.

Initial degradation can cause delayed alarms due to change in dynamics or because of a biased sensor. Nuisance of false alarms due to degrading sensors is important when it comes to the plant operator situation. In power plants, clogging of pressure sensors and mass flow sensors is an effect of fouling. Erosion is due to streaks of flue gas transporting bed sand or fuel particles hitting boiler walls or other components and thus causing damage. Heat exchangers are also exposed to erosion, often in combination with fouling. Some parts of the boiler are exposed to corrosion and are protected, but corrosion can appear on surfaces not designed for aggressive chemicals, due to e.g. fuel mix properties. All the mentioned phenomena can cause plant shutdown, and need to be monitored to take preventive actions and plan maintenance. It is difficult to measure when this degree of degradation moves into the field of loss of control, especially when the degradation is slow and controllability is lost within the alarm limits.

There are tools for detection of degradation using statistics, but they need to be performed off-line by an engineer. To solve the problem by increased maintenance is costly. Sensors showing the same value over time can be faulty even though quality measures like standard deviation and drift trends show excellent values. This kind of degradation is causing loss of control and prevent maintenance because no fault is indicated, the fault remain hidden until extensive data mining have been performed to find such sensors.

Optimisation, financial reports, environmental reports and other top applications using measurements also suffer from degrading sensors. Examples of how degrading sensors and process can affect a process can, for example, be found in reports on recovery boilers used in pulp and paper industry [Doy72] and [Kle98]. The Black Liquor Recovery Boiler Advisory Committee has documented 156 explosions and 450 near-miss incidents in the last 35 years. Following quote from Lefebvre and Santyr [Lef92] comment the report. “In addition to equipment damage, some of the more severe explosions have resulted in injury or even death of operating personnel. There have also been several hundred emergency shutdowns where the fear of an explosion has led to a forced outage. The frequency of explosions has remained relatively constant over the years, and the problem cannot be considered solved”. More examples are found in [Doy72] and [Kle98].

1.4. Literature review

In data treatment several research fields are involved such as; graph theory, optimisation, statistics and modelling. The mathematics needs to be complemented by engineering knowledge to achieve an efficient system for data treatment. An efficient system for data

treatment thus demands many disciplines to interact. Chemical, mineral, control and nuclear power engineers close to applications in respectively area have performed much of the research. The three key problems; detection, isolation and data reconciliation have been developed more or less in parallel. Reviews have been written from the artificial intelligence perspective [Gri92], and more specific on steady state data reconciliation, [Cro96]. Bagajewicz wrote a short review [Bag00] grasping the fundamentals of model-based data treatment and sensor location. Recently Venkatsubramanian published a three-part review on process fault detection and diagnosis, [Ven03a], [Ven03b] and [Ven03c]. The key problems are further reviewed with a timeline in the following sub-sections.

1.4.1. Detection

The first methods for detection were statistical tests, later Kalman filter (KF) [Kal60], principal component analysis (PCA) [Ton94], artificial neural networks (ANN) [Ter93] and Bayesian networks (BN) [Pea88]. These were applied for detection purposes. Detection is used for identifying if there are one or more errors in data, but nothing can be said about the number or location of the errors in these tests. A time-line can be drawn for the development of such tests:

1963	A batch test called global test (GT) is proposed by Reilly and Carpani [Rei63], at a conference. The test is a statistical hypothesis test for normal distribution of random errors in the measurements. At the same conference the nodal test (NT) was proposed by Reilly and Carpani. NT is a hypothesis test on each equation residual.
1982	Mah and Tamhane proposed a hypothesis test, the measurement test (MT) [Mah82], based on the adjustments made in the data reconciliation. The adjustments are tested for normal distribution.
1987	Tamhane, [Tam88a] and [Tam88b], proposed a gross error detection method based on Bayesian networks (BN).
1993	Terry and Himmelblau, [Ter93], proposed an ANN for steady-state gross error detection in measurement on a heat exchanger.
1994	Tong and Crowe, [Ton94], proposed a principal component test for detection of gross errors.
2001	Marshall proposed a gross error detection method based on analysis of the residual from a linear programming solution, [Mar01].
2002	Rollins et al., [Rol02], propose a method for batch detection of gross errors in dynamic processes, the dynamical global test (DGT).

1.4.2. Isolation

When a gross error is detected then the next step is to isolate it. A search method is usually applied. Methods for simultaneous detection and isolation have been proposed e.g. general likelihood test (GLR), modified iterative measurement test (MIMT), unbiased estimation technique (UBET), and linear combination technique (LCT). The timeline is partly extracted from the chapters “Treatment of Gross Errors” in [Rom00] and “Multiple gross error identification strategies for steady-state processes” in, [Nar00], described below.

1965	A serial elimination algorithm was first proposed by Ripps, [Rip65], later extended by Nogita 1972, [Nog72], this approach eliminates one measuring element at a time from the set of measurements and each time checks the value of a test function, subsequently choosing the consistent set of data with the minimum variance.
1975	Almasy and Sztano, [Alm75], suggested a procedure based on the statistical properties of the measurements. The method is limited to systems containing a single element with systematic error, and those cases where the ratio of the extreme error to the dispersion of the regular error is not too small.
1976	Mah et al., [Mah76], extensively studied the problem of the identification of the source of gross errors and developed a series of rules based on graph-theoretical results that enhance the effectiveness of the algorithmic search. Exploiting the topology of the process and using available statistical information a test function for each node in the flow graph is developed which is used in an identification scheme by searching along the internal streams.
1981	Romagnoli and Stephanopoulos, [Rom81] and [Rom83], developed a method to sequentially process the constraints in a recursive way to avoid solving the full-scale reconciliation problem, thus speeding up the computations.
1986	Serth and Heenan, [Ser86], developed a test based on the measurement test called the modified iterative measurement test (MIMT), which uses a serial elimination strategy to detect and identify multiple biases in measuring instruments.
1987	The general likelihood ratio was proposed by Narasimhan and Mah, [Nar87]. The method uses a serial compensation strategy (SCS) based on multiple hypothesis tests for any type of error. A simulator must be used to generate signature vectors for each error, later used for identification and isolation.

1992	Rollins and Davis, [Rol92], proposed the unbiased estimation technique (UBET). This approach simultaneously provides unbiased estimates and confidence intervals of process variables when biased measurements and process leaks exist.
1996	Rollins et al., [Rol96], propose the linear combination technique (LCT), which is a method using nodal tests on single nodes and combination of nodes. A set of rules decrease the number of hypothesis tests efficiently.
1996	Dunia et al., [Dun96], propose a principal component isolation algorithm. A sensor validity index is evaluated for each sensor to identify the faulty sensor.
1999	Jiang and Bagajewicz proposed in a series of articles a method for simultaneous identification of leaks and measurement biases and estimation of error magnitudes called the dynamic integral measurement test (DINT), [Jia99].
2002	Abu-el-zeet et al., [Abu02], developed a detection and isolation algorithm based on the clustering technique by [Che98] and the isolation method by [McB95].

1.4.3. Data reconciliation

Detection, isolation and data reconciliation (DR) form an overall data treatment method. DR is the method of correcting measured values to fulfil the process model and restrictions. The parts in data treatment can be more or less interlaced. Here the data reconciliation and estimation part of the methods proposed in literature are reviewed. Many data reconciliation methods use a least squares method to minimise the impact of noise in the measured data. With the improved computer power, different mathematical programming approaches have become interesting for on-line applications. Data treatment methods have developed from handling the basic linear case with all streams measured, to handling; unmeasured streams, nonlinear constraints, dynamic nonlinear systems, multiple gross errors, errors in both process (e.g. leaks) and sensors (e.g. gross errors). With dynamics involved hold-ups in the process and drift in sensors are also considered.

1961	Kuehn and Davidson, [Kue61], is mentioned as the first to propose data reconciliation. They analytically solved a linear data reconciliation problem where the process was in steady-state with all variables measured and no gross errors present.
------	---

1969-1976	Václavek et al analysed data reconciliation in a series of articles [Vac69], [Vac72], [Vat73] and [Vac76], and worked out the basic principles in data reconciliation of exploiting the process topology to reduce the general problem to a smaller one only incorporating measured variables. They also put forth the concepts of observability and redundancy.
1975	Mah et al. sorted out the relationship between algebra and graph theory in a paper in 1975, [Mah75]. In the same paper gross errors and process leaks were treated for a simulated refinery process, showing that the data reconciliation improved measurement accuracy. In order to reduce the size of the general data reconciliation problem, Mah et al. proposed a method for separating the problem into two sub-problems.
1980	Knepper and Gorman put the focus on nonlinear data reconciliation. Their approach was to use successive linearization followed by solving a linear programming (LP) problem, [Kne80].
1983	Crowe et al., [Cro83], separated measured and unmeasured variables with a projection matrix based on the QR-factorisation algorithm. Classification is an important step in data reconciliation to reduce the size of the general estimation problem.
1986	Crowe et al., [Cro86], expanded the projection matrix method to handle bilinear equation systems.
1992	Liebman et al., [Lie92], applied nonlinear programming technique (NLP) with a moving horizon window on a nonlinear and dynamic process.
1993	Terry and Himmelblau, [Ter93], proposed a method that via ANN performed data reconciliation on the sensors monitoring a heat exchanger.
2000	Soderstrom et al., [Sod00], presented a large-scale industrial application of a dynamic data reconciliation strategy based on the work of Liebman et al, [Lie92]. NLP was used for solving the data reconciliation and estimation problem.
2001	Soderstrom et al., [Sod01], proposed a method for simultaneous gross error identification and reconciliation based on a mixed integer linear programming (MILP) problem formulation.
2003	Bagajewicz and Cabrera, [Bag03], studied a gas pipeline system and proposed an iterative process to solve the data reconciliation problem to account for the error of the approximate mass transfer model equations.

2 Data Treatment

Data treatment is here the layer of algorithms between raw sensor measurement and applications using measurement values. Recall Figure 1 of the data treatment system. Data treatment methods presented here, all use model-based approaches and therefore it is natural to first introduce the process model and the sensors used to measure process state in the real process. Classification of sensors is performed to determine what parts of the system that can be used for detection, isolation, data reconciliation and estimation. Detection is not an essential part of a data treatment system, but it prevents unnecessary search for errors by only triggering isolation when an error is detected. Isolation is the search for one or more detected errors. Data reconciliation is the task of reducing measurement noise by using a process model to compute estimates of the measured value. Finally, estimation of not measured values and computation of error magnitude is the last data treatment step before presenting the data to the top applications.

2.1. *Process models*

The models concerned here are abstract models built of data and knowledge. The model is mathematical and incorporate relation between the variables. It captures the important features. The motives to build models are described by Williams [Wil99], in the following three quotes:

- “The actual exercise of building a model often reveals relationships which were not apparent to many people. As a result a greater understanding is achieved of the object being modelled.”
- “Having built a model it is usually possible to analyse it mathematically to help suggest courses which might not otherwise be apparent.”
- “Experimentation is possible with a model whereas it is often not possible or desirable to experiment with the object being modelled. ...”

The second quote is the most important for the purposes of detection, isolation and data reconciliation. The first quote addresses the importance of a model for diagnosis and the third the possibility to do experiments without risk. Process models are important for many tasks in a thermal process. The models treated here can be used for on-line purposes e.g. fault diagnosis, gross error detection and isolation, process control, state estimation and many other applications. Here mathematical models are intended for state estimation and diagnosis of the process and the sensors it models. In both applications it is important to have a dialogue with the process plant engineer when reconciling or diagnosing with help of a model, because the

model is just a mirror of a sub-set of the real process variables and their relations, nothing more.

How to choose what kind of process model that is suitable for an application is set by what the model is supposed to be used for, what process data are available and by economical constraints. A first principles model is typically more engineering intensive to build and is more expensive than to create a data-driven model. Data-driven model needs a data set for building and also a validation data set.

Analytical models built on first principles are powerful, but demand much time to build. On the other hand, the model is reusable in another way than what is the case for a purely data-driven model, e.g. if the process is retrofitted. An example of a data-driven model is the Bayesian network model; it can be both data-driven and knowledge-based depending on how much of the network and probability tables that are derived from data and from process operator and engineer knowledge.

There are two main principles of how to build a model; analytical models based on relations between variables and parameters observable in physical reality, for example the heat transfer coefficient. Data-driven methods extract relations from a data set, where the model is a fit between given input and output. A parameter in the data-driven model is seldom coupled to physical reality.

Different amount of information is needed for different purposes in plant operation. For alarm purposes, the absolute value of the sensor is enough to check if a threshold value is violated or not. This is done momentarily. If a longer time span for decision or preventive action is needed due to process dynamics and risk, then the information for momentarily triggered alarms is not sufficient. A database for storage of measurement data must then be created. Compressed information such as standard deviations, average values, trends, and so on can be exploited from the historian. If the data is further exploited then data-driven models can be extracted from process measurement data. These models can be used for predictions and early warnings, [Wei01].

The problem of degradation mentioned in the problem section cannot be eliminated for data-driven process models built from measurement data, because there cannot be certainty that data do not contain gross errors or other faults generated from degraded sensors. Updating the model regularly and calibration of sensors used for model building can help to handle the problem. A data-driven model is not valid if one of its inputs is removed. The same problem does not exist for an analytical model, if the variables are classified to determine observability in the model each time a sensor is malfunctioning.

2.2. Sensors

Sensors are needed for control and monitoring. In a heat and power plant, typically mass flows, temperatures and pressures are measured throughout the steam, flue gas and water systems. The most common temperature sensors are working with two different principles; one is the change of resistance in a material, proportional to the change of temperature. These sensors are used for measuring low to medium high temperatures. The second principle of thermocoupling is the change of galvanic potential between two different materials proportional to temperature. Thermocouples can work with high temperatures and are stable over time. Measuring gas temperatures inside the boiler at about 1400 K, (in a solid fuel boiler), is not possible with a physical body in the flue gas stream. Measurement methods such as laser technology and heat cameras are used at such high temperatures.

Pressure is measured as absolute pressure with vacuum as zero on the measurement scale or as relative pressure with atmospheric pressure as zero. The principle of a sensor body that change potential or capacitance with pressure is used in all pressure ranges. Single sensors are installed to measure absolute or relative pressure for control or alarm.

Mass flow can be measured by using two pressure sensors on each side of a contraction to measure the pressure drop. The correlation between flow and pressure drop is known for different geometries, thus the pressure drop can be used for computing mass or volume flow. Other techniques for mass flow measurement are ultra-sound, gravimetric and calorimetric method, [Ols94]. For a boiler the flow composition measurements are used for combustion control and emission reports. The oxygen content is measured to control combustion airflows. Emission reports are based on components such as; Nitrogen oxides, Sulphur dioxide, Carbon monoxide and particulates.

Sensor placement and the sensor accuracy are fundamental for performance of any data treatment system based on sensor readings, but this is out of the scope of this thesis. However, this topic needs some attention for pedagogic reasons. The problem of sensor location has been investigated by Bagajewicz, [Bag97]. An additional handful of researchers have also addressed this topic, [Ali96] and [Rao99]. Positioning the sensors can many times be more important for overall accuracy than the measured standard deviation of a specific sensor in the test bench.

Placement of a sensor can be optimised if parameters as observability, standard deviation, accuracy and so on can be coupled to costs. Such an optimisation gives the minimum instrumentation for producing a product of a given quality. However, this approach is not

easily implemented due to that the accuracy of a sensor is dependent on placement and the environment the sensor is working in.

Trends of increased automation and function-based maintenance raise the demand for automatic diagnose and decision support to the plant operators and maintenance staff. Plants in general are not equipped with enough instrumentation to give redundancy in sensors, which is the theoretical base for analytical diagnosing of an error in a sensor. Installing instrumentation enough for diagnosing sensor errors make it possible to estimate the true process state. It opens the field for the methods described earlier, and also new top applications like advanced control and functional maintenance. The cost for extra instrumentation and maintenance are weighted against benefits from automated diagnosis, early alarms and improved overall control of the process.

A sensor can deviate from the calibrated curve due to many factors. Some of these factors can be corrected by performing new calibration. The alternative is to correct these errors by some tuning parameter or to implement a data treatment system that provides on-line correction by data reconciliation or estimation of the sensor reading.

2.3. Classification of sensors

For complex processes the number of sensors is large. An attempt to set up the general estimation problem for such a process will result in very large equation systems. Classification of sensors was introduced to reduce the general estimation problem. Václavek [Vac69] was the first to propose classification. Later other variable classifications methods have been proposed to divide the general data reconciliation problem into two smaller sub-problems and in most cases reduce the problem significantly. Mah and co-workers proposed classification based on graph theory in 1976, [Mah76]. A great step forward was taken when Crowe developed a projection matrix method in 1983, [Cro83]. The unmeasured sensors are eliminated from the constraints equations by premultiplying them with a projection matrix. Later Crowe extended this method to the bilinear case, [Cro86]. Pai and Fisher, [Pai87], developed an algorithm using Crowe's method on nonlinear functions by successive linearization. For the linear case the projection matrix method solve the classification problem in a straightforward manner without iteration. The projection matrix method is solved more efficiently by QR-factorisation; this was first proposed by Swartz, [Swa89], and later by Sánchez and Romagnoli, [San96].

Beside the major methods for classification such as graph theory and projection matrices, Romagnoli and Stephanopoulos solved the classification problem by an output set assignment

algorithm, [Rom81]. Sanchez, et al developed a program package for this in 1992 called PLADAT, [San92]. PLADAT is an algorithm examining what information each equation supplies (redundancy and observability) and thereby dividing the sensors into subsets. Other methods for classification have been developed by Joris and Kalitventzeff 1987, [Jor87], Simpson et al., [Sim92] and Madron, [Mad92].

Observability and redundancy are defined by Narasimhan, [Na00], as: “A variable is said to be observable if it can be estimated by using the measurements and steady-state constraints”, and redundancy: “A measured variable is said to be redundant if it is observable even when its measurement is removed”. As mentioned, one of the goals with classification is to reduce the general estimation problem. The general estimation problem is a composition of two problems. Mah, [Mah90], defined them as:

“The problem of improving the accuracy of process data so that they are consistent with material and energy balances of the system is known as data reconciliation. Simultaneously, there is also the problem of estimating unmeasured process variables, which is known as coaptation.” Thus, Mah has defined coaptation as the problem to be solved after classification, in absence of gross errors. Which sensor belongs to what sub-problem, see Figure 3, is answered by performing classification of the sensors. First the data set is divided into measured and not measured data sets. By classification the measured and not measured data sets are divided in four groups, [Sa92]:

1. Measured	1a. Overdetermined or redundant . These variables may be adjusted in the data reconciliation problem.
	1b. Just measured . Variables not able to be determined from other sensors. Must be taken directly from measurements. Cannot be adjusted.
2. Not measured	2a. Determinable . Variables not measured, but possible to estimate from balance equations in the coaptation problem.
	2b. Indeterminable . Cannot be determined.

2.3.1. Example of classification by graph theory

Graph theory deals with nodes and directed edges connecting the nodes. A signal graph is outlined in Figure 3 for a small mass flow with both measured and unmeasured signals.

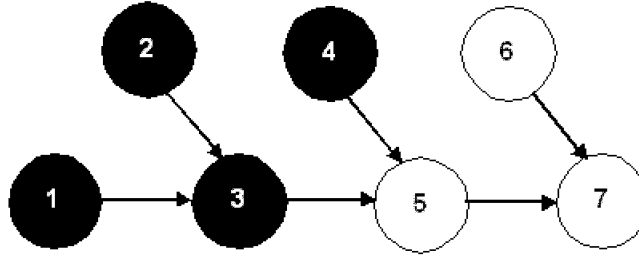


Figure 3. Signal graph. Black coloured signals are measured. White coloured signals are unmeasured.

The signal graph can be used to classify the sensors into the categories presented in the previous section. In graph theory observable and non-observable are denoted as accessible and nonaccessible. Romagnoli, [Rom00], define accessibility as: "We define a node i to be nonaccessible from node a if there is no possibility of reaching node i by starting from a measured node a . Otherwise it is accessible", and determinability as: "We define a node i as determinable if any path going to node i always starts in a measured node".

Look at Figure 3, and use first the definition of accessibility and then determinability. The following results will appear: Nodes 5 and 7 are accessible. Node 6 is nonaccessible. Node 5 is determinable. Nodes 6 and 7 are indeterminable. If we look at the four groups defined earlier then nodes 1, 2, 3 and 4 are measured. Nodes 1, 2 and 3 are over-determined.

2.3.2. Example of classification by matrix projection

The matrix projection method has so far been used on linear and bilinear process models. In this example a small linear process model is considered. The process model in Figure 4 will be used in all the following examples. A fluid is flowing through a small system containing three junctions. The mass balance equations for this process are:

$$s_1 + s_2 - s_3 = 0 \quad \text{Eq. 2.1}$$

$$s_3 + s_4 - s_5 = 0 \quad \text{Eq. 2.2}$$

$$s_5 + s_6 - s_7 = 0 \quad \text{Eq. 2.3}$$

The equations Eq. 2.1 to Eq. 2.3 are normally extracted from a flow-sheet diagram like the one below in Figure 4. The connection between two nodes is called edge in graph theory [Deo74] and arc in mathematical programming, [Win95]. The following text is mainly a subject of matrix projection and graph theory and edge is natural in this context.

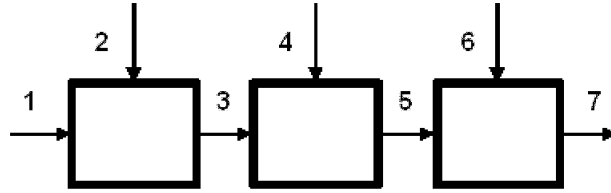


Figure 4. Flow-sheet diagram.

From the flow-sheet for Eq. 2.1 to Eq. 2.3 an linear process model A , which here is the same as the incidence matrix C , can be built, where “1” is for an edge to a node and “-1” is for an edge from a node, and finally “0” is nonexistent edge for a particular node. To couple the matrix A , to a discussion about equations and sensors, then A can be interpreted to contain the three equations Eq. 2.1 to Eq. 2.3 in the rows and the seven mass flow sensors in the columns from left to right.

$$A = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix} \quad \text{Eq. 2.4}$$

If the measured variables are in the vector x and the unmeasured in the vector u , then the mass balance equation is written as:

$$A \cdot \begin{bmatrix} x \\ u \end{bmatrix} = 0 \quad \text{Eq. 2.5}$$

A can be divided into two parts; one containing the measured variables, A_1 , and the other containing the unmeasured variables, A_2 . For the measured variables 1, 2, 3 and 4 we get the matrix A_1 ,

$$A_1 = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{Eq. 2.6}$$

and for the unmeasured variables 5, 6 and 7 we get the matrix A_2 :

$$A_2 = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 1 & -1 \end{bmatrix} \quad \text{Eq. 2.7}$$

The overall mass balance with both measured and unmeasured variables can be written as:

$$A_1 \cdot x + A_2 \cdot u = 0 \quad \text{Eq. 2.8}$$

Matrixes A_1 and A_2 need further treatment to find the sensors that are: redundant, just measured, determinable and indeterminable. A robust method for decoupling the sensors from each other is QR factorisation. This factorisation is similar to the projection matrix proposed by Crowe [Cro83]. QR factorisation of A_2 gives the permuted matrixes Q_u , R_u and the permutation matrix P_u , where index u is for unmeasured variables:

$$A_2 \cdot u = Q_u \cdot R_u \cdot P_u \cdot u \quad \text{Eq. 2.9}$$

If we use that the rank of A_2 (rank is the same as the linearly independent s columns) then we can partition Q_u and R_u into Q_{u1} , Q_{u2} , R_{u1} and R_{u2} , where Q_{u1} and R_{u1} are the s columns of Q_u and R_u .

$$A_2 \cdot u = \begin{bmatrix} \overbrace{Q_{u1}}^s & Q_{u2} \end{bmatrix} \cdot \begin{bmatrix} \overbrace{R_{u1}}^s & R_{u2} \\ 0 & 0 \end{bmatrix} \cdot P_u \cdot u \quad \text{Eq. 2.10}$$

Where the determinable unmeasured variables u_d are the s upper rows of the permuted vector u and the indeterminable u_i are the remaining lower rows of u :

$$u = P_u \cdot \begin{bmatrix} u_d \\ u_i \end{bmatrix} \quad \text{Eq. 2.11}$$

The permuted unmeasured streams are,

$$u = \begin{bmatrix} s_5 \\ s_6 \\ s_7 \end{bmatrix} \quad \text{Eq. 2.12}$$

and u can be divided into u_d and u_i ,

$$u_d = \begin{bmatrix} s_5 \\ s_6 \end{bmatrix} \quad \text{Eq. 2.13}$$

$$u_i = [s_7] \quad \text{Eq. 2.14}$$

thus the unmeasured variables are decoupled into determinable and indeterminable. There is one last operation left; this is to ensure that the determinable variables are not dependent on the indeterminable. After some manipulation of , Eq. 2.5, Eq. 2.8, Eq. 2.9 and Eq. 2.10, we get:

$$u_d = -R_{u1}^{-1} \cdot Q_{u1}^T \cdot A_1 \cdot x - \underbrace{R_{u1}^{-1} \cdot R_{u2}}_{R_{iu}} \cdot u_i \quad \text{Eq. 2.15}$$

and the rows in the product R_{iu} that are not zero indicate that u_d and u_i have a dependency. From the example we get:

$$-R_{u1}^{-1} \cdot R_{u2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ corresponds to } \begin{bmatrix} s_5 \\ s_6 \end{bmatrix} \quad \text{Eq. 2.16}$$

In the steady-state mass flow, stream s_6 is dependent on s_7 , s_6 is therefore also indeterminable. Stream s_5 is independent of s_7 and remain determinable. The fact that s_6 now is indeterminable do not affect the status of s_5 . In Figure 4 we see that s_5 can be determined from the measured streams s_3 and s_4 .

Both Romagnoli [Rom00] and Narasimhan [Nar00] have shown that the projection matrix developed by Crowe [Cro83] is the same as the transpose of the matrix Q_{u2} in Eq. 2.10. Using this knowledge about the projection matrix it is possible to decouple the variables x connected to the measured matrix A_I into redundant x_r and non-redundant x_n :

$$Q_{u2}^T \cdot A_1 \cdot x = \begin{bmatrix} 1 & 1 & -1 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_r \\ x_n \end{bmatrix} \quad \text{Eq. 2.17}$$

x_r and x_n are substituted for their sensor numbering and thus give:

$$(Q_{u2}^T \cdot A_1) \cdot x = \begin{bmatrix} -1 & -1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} \quad \text{Eq. 2.18}$$

The right hand side of Eq. 2.18, from the projection tells us that the first three measured variables; s_1 , s_2 and s_3 are redundant because they are linearly dependent i.e. not zero. s_4 is projected as non-redundant or just measured. Now the classification is completed and the results are:

$$x_r = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \quad \text{Eq. 2.19}$$

$$x_n = [s_4] \quad \text{Eq. 2.20}$$

$$u_d = [s_5] \quad \text{Eq. 2.21}$$

$$u_i = \begin{bmatrix} s_6 \\ s_7 \end{bmatrix} \quad \text{Eq. 2.22}$$

Where x_r include redundant measured variables, x_n non-redundant variables, u_d determinable unmeasured variables and u_i unmeasured indeterminable variables.

The strength of including a classification algorithm in a data treatment system is that all the measured variables and unmeasured variables are monitored on-line for their status. Thus it is easy to determine how much information the present set of functioning sensors can provide. As shown, this can be gained by treatment of a signal graph and linear (or linearized) process equations. How the information from classification is being used will be developed in the following sections.

2.4. *Some basic properties used in statistics*

Noise and fluctuations in the process are common in real applications and probably give sensor readings that are distributed around a mean value. The mean value μ , standard deviation s , and covariance are important properties when doing statistical calculations. The mean value μ_y , for variable y , is calculated as:

$$\mu_y = \frac{\sum_{i=1}^I y_i}{I} \quad \text{Eq. 2.23}$$

where I is the number of samples and index y is indicating measures computed for y . The mean value is used for calculation of the standard deviation,

$$s_y = \sqrt{\frac{\sum_{i=1}^I (y_i - \mu_y)^2}{I - 1}} \quad \text{Eq. 2.24}$$

where the standard deviation s_y is the square root of the variance. Variance is the expected square of the distance from the mean value and the standard deviation is thus the expected distance from the mean value, [Ada96]. Finally a covariance can be computed:

$$\text{cov}(x, y) = \frac{\sum_{i=1}^I (x_i - \mu_x)(y_i - \mu_y)}{I - 1} \quad \text{Eq. 2.25}$$

Covariance shows how variables correlate. In the general case with N variables in columns and M observations in the rows, with corresponding counters n and m , it is convenient to use the vector form for covariance. The samples of all y are ordered in the matrix, Y :

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{MN} \end{bmatrix} \quad \text{Eq. 2.26}$$

The measurement covariance matrix can then be expressed in matrix form as, [Rus00]:

$$\text{cov}(Y) = \frac{1}{M - 1} \cdot Y^T Y \quad \text{Eq. 2.27}$$

In this thesis different hypothesis tests are presented. To perform a hypothesis test, we need a null hypothesis, H_0 and an alternative hypothesis, H_a . For the test we choose significance level α_{sig} that tells the uncertainty in the test. If α_{sig} is 5% then there is 5% chance of rejecting the null hypothesis when it should have been accepted, [Fre00]. This is also called Type I error or in everyday words; false alarm. A Type II error occurs when the alternative hypothesis is rejected although it is true. The test includes a test quantity that is computed, and then evaluated, to decide if the null hypothesis is rejected or not.

Here most of the hypothesis tests are goodness-of-fit tests. In these tests an assumption is made about how the tested variable or variables are distributed. This is exemplified in section 2.6.2.

2.5. Normal and chi-square distributions

In the following sections the sensor readings are assumed to be normally distributed around a mean value. A normal distribution with a mean value $\mu = 0$ and standard deviation $s = 1$ is shown to the left in Figure 5. To the right in Figure 5 is a similar dashed curve, but it has a different mean value $\mu = 2$ and standard deviation $s = 1$.

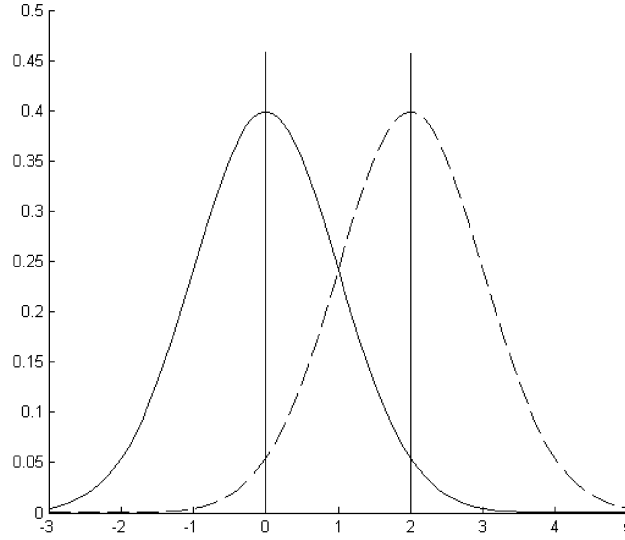


Figure 5. Probability density function curves for normal distribution.
Solid curve has $\mu = 0$ and dashed curve has $\mu = 2$.

The curve above is calculated from the normal distribution probability density function (pdf) [Mat02]:

$$y = \frac{1}{s \cdot \sqrt{2\pi}} \cdot e^{\frac{-(x-\mu)^2}{2s^2}} \quad \text{Eq. 2.28}$$

Where y in this case is the height of the curve. The probability p , for x to be in an interval between x_1 and x_2 is equal to the solution of Eq. 2.29, which is the integral over the pdf in the interval.

$$p = \int_{x_1}^{x_2} \frac{1}{s \cdot \sqrt{2\pi}} \cdot e^{\frac{-(x-\mu)^2}{2s^2}} dx \quad \text{Eq. 2.29}$$

The chi-square distribution is a special case of the gamma distribution, [Mat02]. Later in this section it is used in a hypothesis test for determining if a set of measurement samples contain gross errors. The distributions are families of curves.

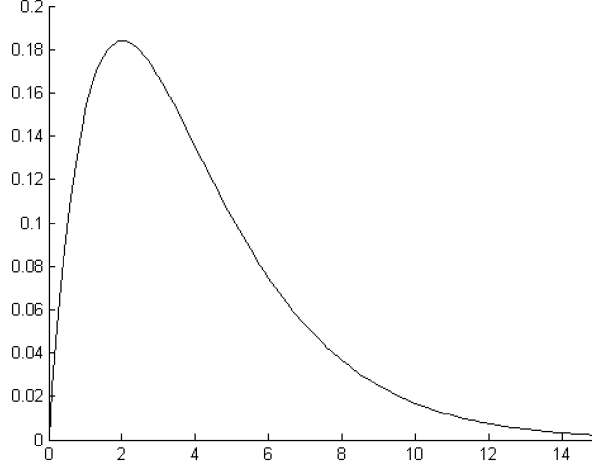


Figure 6. A curve from the chi-square pdf.

The chi-square pdf is computationally more expensive to evaluate than the normal distribution pdf. The chi-square pdf is:

$$P_k(x) = \frac{\left(\frac{1}{2}\right)^{k/2}}{\Gamma(k/2)} \cdot x^{k/2-1} \cdot e^{-x/2} \quad \text{for } x > 0 \quad \text{Eq. 2.30}$$

where Γ is the gamma function, k are the degrees of freedom and x is the sample value. Analysis shows that the gamma function of a half integer is the same as a chi-square distribution, [Hog78]. The gamma function is expressed as:

$$\Gamma\left(\frac{k}{2}\right) = \int_0^{\infty} x^{k/2-1} \cdot e^{-x} dx \quad \text{Eq. 2.31}$$

There are tables with values for the chi-square distribution for each degree of freedom and value of the significance level α_{sig} . The chi-square probability distribution function is a family of curves with different appearance due to degrees of freedom. Degrees of freedom are explained by Weiss, [Wei04], in the following quote, ‘The number of degrees of freedom in a problem, distribution, etc., is the number of parameters which may be independently varied’. The statistical degrees of freedom are often called redundancy in the field of data reconciliation. Here degree of freedom r is defined by a short expression, [Rom00]:

$$r = l - g \quad \text{Eq. 2.32}$$

Where l is the number of measured variables and g is the number of estimated variables.

2.6. Detection of gross errors

Presence of gross errors in measurement data decrease efficiency of data reconciliation, [Rom00] and [Nar00]. Thus gross errors need to be removed before data reconciliation. Detection is the first step to perform to be able to remove a gross error. When talking about detection it is also important to discuss what to detect. According to Kesavan [Kes97] the types of faults that occur in a chemical process include process parameter changes, disturbance parameter changes, actuator problems and sensor problems. This is applicable to the heat and power plant industry as well. There is no strict definition of gross error, but the one used here was stated by Veverka and Madron in 1997, [Mad97]:

“Statistically, a gross error is an error whose occurrence as realization of a random variable is highly unlikely”.

This definition leaves everything that the data reconciliation cannot handle to the detection and isolation algorithms to solve. A model of measurement is introduced to define what it is the tests aims to detect in the sensors. From the data reconciliation point of view the measured signal can be divided into three main components; the estimated true state of the process x , the gross error δ , and the random noise e . From these components the measured signal, y , can be modelled as:

$$y = x + \delta + e$$

Eq. 2.33

Noise e is assumed to be normal distributed and have an expected mean value equal to zero. The gross error δ is assumed to be independent of signal magnitude. If the gross error can be detected, then an isolation algorithm is triggered and if possible the gross error is isolated and removed. Here the focus is on statistical hypothesis tests.

The development began in the 1960s. Ripps proposed a method based on measurement elimination in 1965, [Rip65]. Reilly and Carpani proposed the statistical hypothesis batch test called global test at a conference, [Rei63]. This test was followed by a number of statistical tests: general likelihood ratio test [Wil74], measurement test [Alm75], and nodal test [Mah76]. The mentioned tests were further developed during the 1980s. Besides the statistical hypothesis tests for gross error detection, other techniques have been proposed, for example; artificial neural networks, Bayesian networks, principal component analysis and Kalman filter [Mak95]. Recent developed methods [Rol00] are in the field of multiple gross errors detection in dynamic processes.

2.6.1. Tests for detection

In this section hypothesis tests are presented e.g. global test, nodal test, measurement test and general likelihood ratio test. Gross error detection is used to activate the computational expensive search for gross error only when there are detectable gross errors present. Goodness-of-fit tests have shown to be quite good for the task, [Sun03]. Assumptions are independent variables and normal distributed noise in the variables with zero mean. The tests aims to detect if an observed frequency deviate from an assumed theoretical frequency more than can be expected by pure chance. The first proposed goodness-of-fit test for gross error detection was the global test, [Rei63].

Global Test (GT)

The global test uses the residual r of a linear process model in matrix form A , and the measured variables in the vector y . If the process model is interpreted as constraints, then the residual represents the violation of the constraints expressed in a vector. The residual vector can be used to compute a test quantity for a collective test for gross errors. If there are no gross errors present then the residual follows a normal distribution with zero mean. The residual of the process model and the measurements is computed as:

$$r = A \cdot y \quad \text{Eq. 2.34}$$

To form a test quantity for the global test, first the covariance needs to be computed. Below, the covariance matrix V is calculated from the residual vector r , and the number of variables n :

$$V = \frac{1}{n-1} r^T r \quad \text{Eq. 2.35}$$

If the measurement error covariance matrix Σ is known, then V can be calculated according to:

$$V = A \cdot \Sigma \cdot A^T \quad \text{Eq. 2.36}$$

There are many difficulties to obtain the measurement error covariance matrix and this is investigated in the book by Narasimhan, [Nar00], pages 77-81. Also Morad et al., [Mor99], have investigated this. There are doubts if the more sophisticated methods to get the error measurement covariance matrix are better than the straightforward direct method. The direct method is often used to obtain the measurement error covariance matrix directly from process data, but it can be discussed whether this is a good method since the data used for computation can be contaminated with gross errors. The variables are weighted according to

the inverse of the covariance matrix, V , and then form the global test statistic, γ , that follows a chi-square-distribution with degrees of freedom equal to the rank of A .

$$\gamma = r^T \cdot V^{-1} \cdot r \quad \text{Eq. 2.37}$$

The variance is here a measure of how reliable a measurement is for a process in steady state. It can be discussed if it is good because an increase or decrease of variance for a sensor in good shape can in both cases mean that the sensor is degrading or affected by e.g. fouling.

The null hypothesis is that there are no gross errors present in the measurement data. This can be falsified by the global test. Gross error is probably present if γ exceeds the theoretical value of the chi-square-distribution at the chosen level of significance α_{sig} and ν degrees of freedom, this conclusion relies on the assumption that each sensor signal got a normal distribution. If the null hypothesis is not rejected, data is probably free of gross errors or the gross errors have eliminated each other and are thereby concealed. The latter is improbable if there are many measurements and high interconnection between the sensors. In a mass balance model, large magnitude in a signal with high standard deviation can conceal smaller magnitudes and cancel the possibility to find gross errors in them.

Nodal Test (NT)

A total of M constraints are tested for gross error. The residual r_i of each constraint follows a normal standard distribution. Because we have M tests instead of one batch test as in the global test, the probability increase that the null hypothesis is rejected. The limit α_{sig} needs to be modified to handle multiple hypothesis tests, without changing the significance level. The modified α_{sig} called β can be computed as, [Rol92]:

$$\beta = \frac{\alpha_{sig}}{M} \quad \text{Eq. 2.38}$$

The null hypothesis of the nodal test is: Data is normal distributed and the test quantity, γ , follows a normal distribution. The alternative hypothesis is: Data is not normally distributed. The test statistics to be evaluated is,

$$\gamma = \frac{|r_i|}{\sqrt{V_{ii}}} \quad \text{Eq. 2.39}$$

where r_i is the residual of constraint i , and V_{ii} is the corresponding covariance value in the diagonal of V . The critical value $Z_{1-\beta/2}$ is determined from a standard normal distribution table

with input of β . If $\gamma \geq Z_{1-\beta/2}$ then the null hypothesis is rejected and the residual for constraint m is probably containing a gross error.

Measurement Test (MT)

The measurement test is a test on the computed adjustments. This test is dependent on a data reconciliation algorithm that conserves the information about the normal distribution in the variables into the adjustments. Adjustments are computed after the data reconciliation step as the difference between the measured value y and the data reconciled estimates x . Similar to the nodal test a test quantity is computed and tested for normal distribution.

General Likelihood Ratio (GLR)

This approach of gross error detection needs a model of the system. Leaks can be detected by the method. A simulator is fed with data containing gross errors and leaks. The response from the system is recorded into signature vectors. The GLR test the hypothesis, that the measured value is contaminated with gross error or leak is less likely, than the measured values being just normally distributed for the given measurement vector with zero mean. The test statistic γ is the ratio between the probability of the present null hypothesis and the alternative hypothesis [Nar00]:

$$\gamma = \sup \frac{\Pr\{r|H_1\}}{\Pr\{r|H_0\}} \quad \text{Eq. 2.40}$$

Here \sup is the calculation of the supremum over all possible values of the parameters in the hypothesis. Calculation of Eq. 2.40 is made for all signature vectors. The test uses multiple tests, compare the nodal test, and it also needs modification of the significance level α to β as in Eq. 2.38 before comparing the test statistics to a chi-square distribution.

2.6.2. Example of gross error detection by global test

In this example of gross error detection, the process scheme in Figure 4 is used. The example got the following scenario: all mass flows are measured. One of the mass flow meters has been partially clogged/fouled and show lower mass flow than its true value.

A vector is created containing noise to be reconciled and gross error to be detected. The true values for measurement vector y are: [2 2 4 1 5 3 8] (kg/s). For simulation a 5% relative noise is added to all measurements, and we get the (simulated) measured values, $y_{meas} = [1.9398 \quad 1.9031 \quad 4.0987 \quad 0.9945 \quad 5.2159 \quad 2.9898 \quad 7.9349]$ (kg/s). Due to

clogging/fouling a gross error of -3 kg/s is introduced in sensor 5. Adding this error to y_{meas} gives the vector $y_{meas,GE} = [1.9398 \quad 1.9031 \quad 4.0987 \quad 0.9945 \quad 2.2159 \quad 2.9898 \quad 7.9349]$. Assume the measured values are normally distributed, thus the residual of a linear equation system is also normally distributed. Measurement covariance matrix is here an identity matrix. The graph is transformed to the incidence matrix A , see Eq. 2.4. Now the global test is used to detect gross error in the measured values y_{meas} and $y_{meas,GE}$. The global test quantity is calculated for measured vector according to Eq. 2.37. The null hypothesis is: the residual follows a chi-square distribution and thus the measurements do not contain gross error if $\gamma \leq Z_{critical,v}$, where $Z_{critical,v}$ is the theoretic threshold for chi-square distribution with v degrees of freedom and significance level of 95%. $Z_{critical,v}$ is tabulated in literature for degrees of freedom and significance level, [Fre00]. In this case degrees of freedom equal to 3, and significance level 95%. With these inputs we get the critical test quantity value:

$$Z_{critical,v} = 7.654 \quad \text{Eq. 2.41}$$

The test quantity γ_1 for the measured vector y_{meas} containing only measurement noise may be formulated as below after substituting the residual r in Eq. 2.37 with Eq. 2.34, the resulting expression is:

$$\gamma_1 = (A \cdot y_{meas})^T \cdot V^{-1} \cdot (A \cdot y_{meas}) \quad \text{Eq. 2.42}$$

giving:

$$\gamma_1 = 0.1539$$

For the vector $y_{meas,GE}$ containing both measurement noise and gross error of -3 kg/s in sensor 5, the test quantity γ_2 is calculated as in Eq. 2.42:

$$\gamma_2 = (A \cdot y_{meas,GE})^T \cdot V^{-1} \cdot (A \cdot y_{meas,GE}) \quad \text{Eq. 2.43}$$

giving:

$$\gamma_2 = 15.793$$

The test quantity γ_1 computed for vector y_{meas} , is less than the critical value $Z_{critical,v}$, thus the null hypothesis is not rejected and the measured vector y_{meas} is not considered containing gross error. This implies that deviation in measured values is explained by random noise in the sensors. The second test quantity γ_2 computed from the vector $y_{meas,GE}$ shows to be larger

than the critical value $Z_{critical,v}$. Thereby the null hypothesis is rejected implying that $y_{meas,GE}$ contain one or more gross errors.

2.7. Isolation

The task of the isolation algorithm is to search for errors. If a gross error or other fault is found then the error either eliminated or corrected. Isolation of errors may be seen as a protection for the following data reconciliation algorithm, see Figure 1. Most data reconciliation algorithms cannot handle errors in data. Ripps proposed isolation of gross errors in 1963, [Rip63]. The first isolation algorithms were based on graph theory. Data driven techniques such as artificial neural networks (ANN), [Arr03], Bayesian networks (BN) [Wei01] and Kalman filters (KF) [Mak95] have also been proposed for isolation. The growth in computational power has also opened up for optimisation by mixed integer programming (MIP). MIP gives rise to very large optimisation problems even for a moderate sized sensor network when multiple simultaneous errors are considered. A common way to handle this is to make a single fault assumption or a “few faults” assumption. In a work by Gatzke [Gat01], two simultaneous errors was the limit of maximum simultaneous errors, thus the search tree is pruned to a manageable size.

When to use multiple and single fault assumptions is depending on how often errors occur, the length of sampling time, complexity of the system and its equations, need for isolation of multiple errors, and of course the cost compared to the benefit of the system. The cost incur from extra sensors and their maintenance, hardware for the software, installation cost and implementation of the software in the system. The benefits of isolation are; shorter time to find sensor and process faults, and a possibility to estimate the size of the fault. If the fault is both isolated and known in size then the plant can be in operation without need to shut down until next maintenance is scheduled, thus expensive unscheduled shutdowns can be avoided.

The idea is to perform single gross error isolation for each sensor and then eliminates the sensor containing a gross error until all sensors have been tested for gross error. How the algorithm can be performed is shown in Figure 7. The benefit is that this method only increases proportional to the number of sensors in computational work. The drawback is that it is not a completely robust method, because the sequence of elimination of the errors has an effect on which the next detected and isolated sensor will be.

If the signals are consequent by time and increasing/decreasing, the probability for a real fault increases, while a shifting signal may just be noise. See 2.7.4 for an example. An algorithm built to isolate more than one gross error has a search tree that grows exponentially

with the number of simultaneous gross errors the algorithm needs to isolate. Different methods have been proposed to decrease the exponential growth. For example the linear combination technique (LCT), [Rol96], use conclusions about the interconnections in the graph to prune the search tree. Adjustability tells if a variable can be adjusted or not, [Sun03]. This information can be used for eliminating sensors that are not adjustable from the search tree.

2.7.1. Isolation of single gross error based on statistical hypothesis test

In the data set T contains t sensors. A test quantity Q is computed for each sensor. The criterion for the test quantity is checked for each sensor and on basis of that a sensor t is picked containing gross error. Indication of gross error is the largest value of all Q .

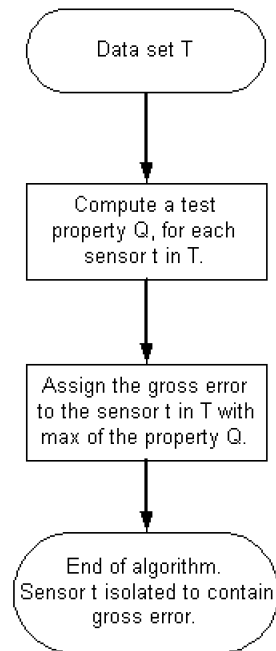


Figure 7. Isolation based on statistical tests for single gross error.

Isolation of one gross error in a sensor is performed by a straightforward method sufficient for small applications. When the sensors increase in number, the possibility of more than one simultaneous error increase and multiple errors must be assumed. The straightforward search

method is common for data-driven methods that seldom use loops in the algorithm. Multiple gross error search methods based on statistical tests can use the search method for single gross error, but then it is incorporated in a loop.

2.7.2. Isolation of multiple gross errors based on statistical hypothesis test

Here the sequence for search of single gross error is complemented with a detection of gross error test. When no more gross errors are detected the loop terminates. Each gross error found is put into the data set S and is eliminated from the data set T. Thus, all sensors containing gross error is in the data set S when the loop terminates and the sensors without gross error are in data set T.

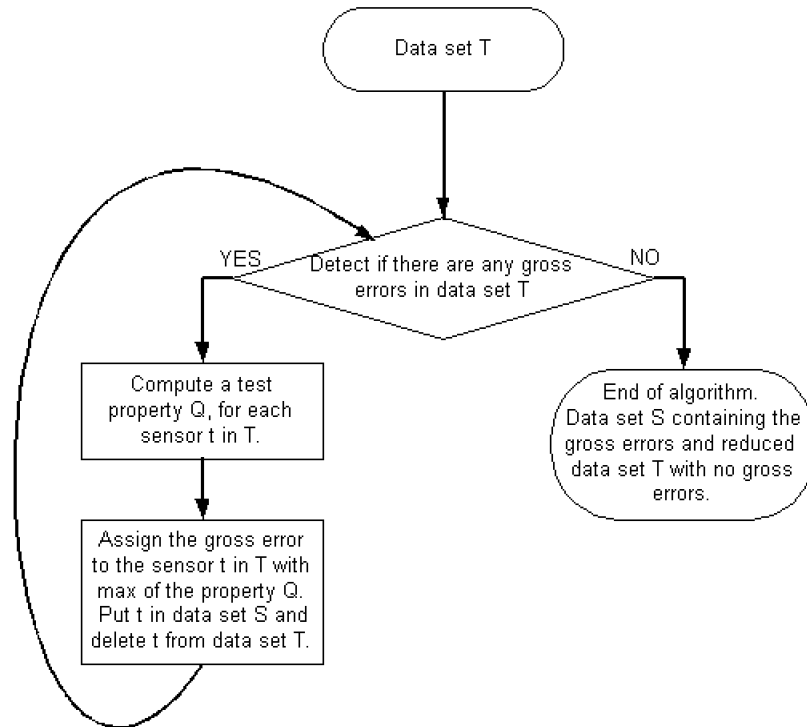


Figure 8. Isolation based on statistical tests for multiple gross errors.

The above problem can also be solved in one problem formulation using either linear or nonlinear mixed integer programming (MILP and MINLP). Often nonlinear programming (NLP) is used as a reference method. Solving a NLP-problem is time consuming and is not yet thought of as usable in on-line applications.

2.7.3. Isolation of gross error by solving a sequence of nonlinear programming problems

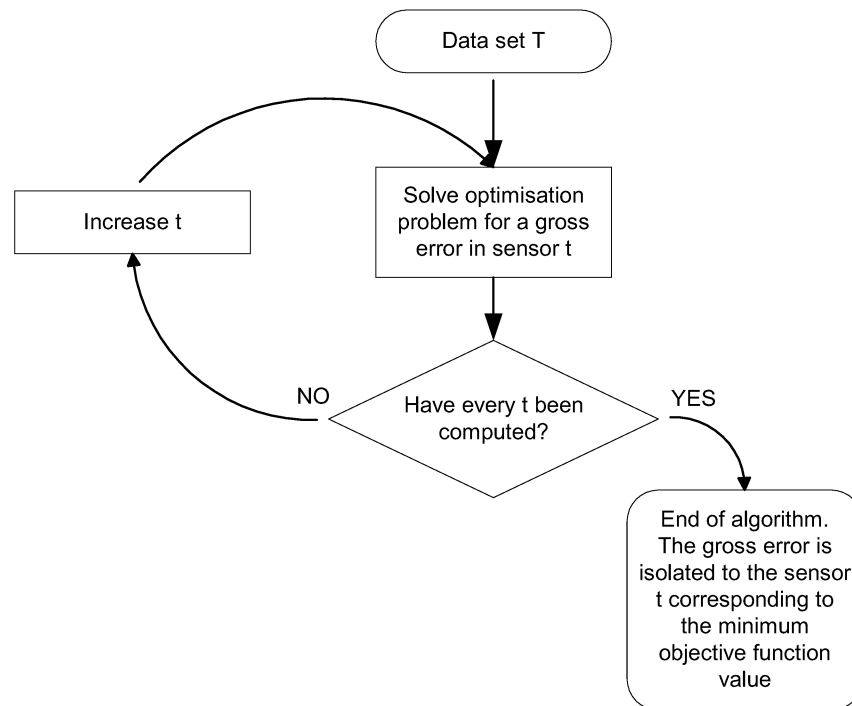


Figure 9. Isolation based on solving a sequence of nonlinear programming problems.

Solving a sequence of NLP problems give an objective function value for every sensor in the process. The general problem set up for NLP problems and the ability to handle time window, give possibility to incorporate different types of faults. For example the time window can be used for estimating fault development or dynamic faults like dampening, which is not possible to detect in a steady-state solution.

2.7.4. Example of isolation by solving a sequence of NLP problems

We continue on the previous example, with the given matrix A , and the measured vector $y_{meas,GE} = [1.9398 \ 1.9031 \ 4.0987 \ 0.9945 \ 2.2159 \ 2.9898 \ 7.9349]$ containing a gross error of -3 kg/s in sensor number 5. Suppose a gross error was detected. The isolation algorithm is activated and the task is to isolate a single gross error. We choose the method of solving a sequence of NLP problems for isolation:

$$\min_{\delta, x} \left(\sum_{i=1}^I \sum_{n=1}^N (x_{i,n} - y_{i,n})^2 + \mu \sum_{i=1}^I \sum_{m=1}^M \left(\sum_{n=1}^N c_{m,n} x_{i,n} \right)^2 \right)$$

$$s.t. \quad y_{i,n} = x_{i,n}, \quad n \neq n', \quad i = 1, \dots, I$$

$$y_{i,n} = x_{i,n} + \delta, \quad n = n', \quad i = 1, \dots, I$$
Eq. 2.44

Where y is the sensor reading, x is the estimated value of a sensor reading, c is an element in the incidence matrix A , μ is the penalty function multiplier, gross error δ , and M , N and I are total number of rows, columns and samples, respectively. The sensor evaluated in each solution is n' , the other $N-1$ sensors are not changed. To be able to exploit the features of the time window approach, 10 samples are simulated. The 10 samples shown in Table 1 below are used as input to the NLP problem in Eq. 2.44.

<i>Sensor reading [kg/s]</i>						
<i>S1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>S5</i>	<i>S6</i>	<i>S7</i>
1.9934	2.0062	4.0233	0.9807	2.0996	3.0382	8.1164
2.0315	2.0170	3.9402	0.9887	2.0316	3.0055	7.8238
1.9589	1.9771	3.9818	0.9987	2.1022	3.0144	7.9316
1.9978	2.0097	3.9323	1.0165	2.1140	3.0143	7.8115
2.0312	2.0110	4.0403	0.9796	1.9812	2.9813	7.8665
2.0333	2.0339	3.9903	1.0228	1.9118	3.0555	8.1078
1.9944	2.0121	4.0903	1.0070	1.9368	2.9779	7.8751
1.9991	1.9909	3.9927	1.0055	1.8928	2.9721	8.0434
1.9675	2.0121	3.9492	1.0044	2.0015	2.9947	8.0166
2.0442	1.9842	3.9804	0.9904	1.9779	2.9679	7.9577

Table 1. Sensor readings with gross error in sensor 5 equal to -3 kg/s in each sample.

Results from solving the sequence of problems Eq. 2.44 with the samples in Table 1 are shown in Table 2 below.

<i>Sensor</i>	<i>S1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>S5</i>	<i>S6</i>	<i>S7</i>
<i>Estimated gross error [kg/s]</i>	0.018	0.018	1.484	2.986	-2.967	-2.948	2.948
<i>Objective function value</i>	1765	1765	1324	873	4	896	896
					<i>Min!</i>		

Table 2. Result from solving a sequence of NLP problem for each sensor to isolate the gross error.

In Table 2 the minimum function value from solving Eq. 2.44 is found for sensor 5, thus a GE in sensor 5 with the size of -2.97 kg/s is the best solution that fulfils the model. This result complies very well with the simulated GE of -3 kg/s.

2.8. Data Reconciliation

Data reconciliation (DR) is a special case of the general estimation problem. DR is the problem of having measurements in an over-determined system satisfying all process constraints. Unmeasured variables are either estimated simultaneously with DR in a general estimation problem or as a separate step after the DR problem is solved. Solving DR and estimation simultaneously is in some literature called data coaptation. Data reconciliation has been reported in applications for mineral industry processes [Sim91], chemical production plants [Abu03], and nuclear power plants, [Sun03]. The most active application field for research in data reconciliation is in chemical engineering. The first method of data reconciliation for a chemical process was proposed in 1961, [Kue61]. The methods involved linear system equations under the assumption that all process variables were measured and absence of gross errors. The described problem can be solved analytically by least squares methods. Complementary methods as gross error detection and isolation where developed parallel to DR in order to handle real DR application problems. Graph theory was connected to the DR problem formulation by Mah et al., [Mah75]. The assumption of unmeasured variables was not efficiently solved until Crowe [Cro83], presented the projection matrix method. Swartz [Swa89], showed that the projection matrix operation could be solved by the robust QR-factorisation.

Nonlinear data reconciliation was first proposed to be solved by successive linearization by Knepper and Gorman, [Kne80]. Liebman et al., [Lie91], showed that nonlinear programming methods gave improvements in accuracy, but had long computation time. Liebman et al., [Lie92], later proposed a time window approach for dynamic data reconciliation. Russo and Young implemented a dynamic DR algorithm in 1999, [Rus99]. Recently Soderstrom et al., and Gatzke et al., reported industrial implementations in chemical plants, [Sod00] and [Gat02], using nonlinear dynamic DR strategies.

The general data reconciliation problem formulation for linear equations solved for I samples by a least squares method [Rom00] is here formulated as:

$$\begin{aligned} \min_{x,u} \sum_{i=0}^I (y_i - x_i)^2 \\ \text{s.t. } A_1 \cdot x + A_2 \cdot u = 0 \end{aligned} \tag{Eq. 2.45}$$

Where y is the measured state, x is the estimated state, u is the unmeasured state, A_1 is the columns of A with only measured variables and A_2 is columns of A with unmeasured variables. The simultaneous solution of data reconciliation and estimation of unmeasured variables in Eq. 2.45 is in many cases a large task. To reduce the size of the optimisation problem the DR problem can be divided into two sub-problems, [Cro83]. After the division of the total problem, the next task is to solve the over-determined equations. The following sub-problem is to estimate the unmeasured variables with data from the solution of the over-determined equation system. The simplest DR problem with only measured variables presented below is with linear equations with no gross errors present and the process in steady-state:

$$\begin{aligned} \min_x \sum_{i=0}^I (y_i - x_i)^2 \\ \text{s.t. } A_1 \cdot x = 0 \end{aligned} \tag{Eq. 2.46}$$

The linear process model A_1 , containing only measured variables can for example be the conservation laws for a system with mass and energy streams. The following sub-problem solves the linear equation system of unmeasured variables.

$$A_2 \cdot u = 0 \tag{Eq. 2.47}$$

This equation system is solved with any linear equation solution method, for example LU-decomposition or QR-decomposition. The sensors have different measurement error; this can be implemented in the problem formulation as a weight, w , on each sensor. How this weight is determined does not follow any standard procedure. In [Sod00] an estimated accuracy of the

sensor is used, which can be modified by the plant operation engineers on the basis of past measurements and their experience and knowledge about the considered instrument. Define a matrix W , where the weights w is put in the diagonal. Now we can formulate the weighted least squares problem,

$$\begin{aligned} \min_{x,u} \quad & \sum_{i=0}^I W_{ii} (y_i - x_i)^2 \\ \text{s.t.} \quad & A_1 \cdot x + A_2 \cdot u = 0 \end{aligned} \tag{Eq. 2.48}$$

where x and u are computed to minimise the objective function value. In a one-sample approach it is difficult to distinguish a large random noise from gross error. Data reconciliation for the present vector y in Eq. 2.48 can be improved by computing a time series of samples. A time window makes more data available for the problem solving and noise can efficiently be smoothed. The cost for the decreased measurement error is increase in computational load, which at least is proportional to the number of time steps in the time window. The DR problem formulation in Eq. 2.48 can be extended with a time window, T :

$$\begin{aligned} \min_{x,u} \quad & \sum_{t=0}^T \sum_{i=0}^I W_{t,ii} (y_{t,i} - x_{t,i})^2 \\ \text{s.t.} \quad & A_1 \cdot x_t + A_2 \cdot u_t = 0, t = 1, \dots, T. \end{aligned} \tag{Eq. 2.49}$$

The solution x and u from Eq. 2.49 is ready to use for applications depending on measured data as, plant optimisation, control purposes, diagnostics, economical and environmental reports.

2.8.1. Example of data reconciliation

We continue on the example used throughout this thesis beginning with classification and ending with data reconciliation. Earlier a gross error was detected in the data set and was isolated to sensor 5. The size of the gross error was estimated to -2.97 kg/s. Use the data set given in Table 1 and correct the measured value in sensor 5 with the estimated gross error. Solve the data reconciliation problem Eq. 2.49 for the corrected data set in Table 3. Observe that there are no unmeasured sensors in this case.

<i>Sensor reading [kg/s]</i>						
<i>S1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>S5</i>	<i>S6</i>	<i>S7</i>
1.9934	2.0062	4.0233	0.9807	5.0696	3.0382	8.1164
2.0315	2.0170	3.9402	0.9887	4.9836	3.0055	7.8238
1.9589	1.9771	3.9818	0.9987	5.0722	3.0144	7.9316
1.9978	2.0097	3.9323	1.0165	5.084	3.0143	7.8115
2.0312	2.0110	4.0403	0.9796	4.9512	2.9813	7.8665
2.0333	2.0339	3.9903	1.0228	4.8818	3.0555	8.1078
1.9944	2.0121	4.0903	1.0070	4.9068	2.9779	7.8751
1.9991	1.9909	3.9927	1.0055	4.8628	2.9721	8.0434
1.9675	2.0121	3.9492	1.0044	4.9715	2.9947	8.0166
2.0442	1.9842	3.9804	0.9904	4.9479	2.9679	7.9577

Table 3. Sensor readings with corrected gross error in sensor 5 equal to -2.97 kg/s.

In this case the least squares problem as in Eq. 2.46 is solved analytically, resulting in the vector $x = [2.0119 \ 2.0247 \ 4.0367 \ 1.0126 \ 5.0493 \ 3.0526 \ 8.1020]$.

Define the residual r , as:

$$r = x - y_{true} \quad \text{Eq. 2.50}$$

where x is the vector of estimated values and y_{true} is the flow vector without noise and gross errors. What effect does the data treatment have on the errors and noise in the measured data? Two-norm of the residual vector r is a measure of how far the true and estimated values are from each other. The two-norm is denoted $\|r\|$. By computing the two-norm, it is possible to follow the contribution of each step in the data treatment.

<i>Step in data treatment</i>	<i>Description of measurement vector</i>	<i>Two-norm of r</i>
Simulation of real measurements by adding noise and gross error	True measurement	0
	True measurement with 5% relative noise	0.1444
	True measurement with gross error	3.0000
Detection	True measurement with gross error and noise	3.0722
Isolation	Corrected measurement after isolation	0.1431
Data reconciliation	Data reconciled measurement after correction of isolated gross error	0.1336

Table 4. Two-norm of the residual in different steps of the data treatment.

To simulate a measurement vector, noise and gross error are added. The gross error was detected by the global test. It was isolated to signal 5 and estimated to -2.97 kg/s in the isolation algorithm. After signal 5 was corrected with 2.97 kg/s (and thereby the gross error was deleted) the two-norm decreased from 3.0722 to 0.1431, a significant improvement. However, the added noise still remains. Using a least squares method for data reconciliation, the noise is reduced and the two-norm is further lowered to 0.1336, which is the remaining error. This value is lower than any of the two-norms for the simulated signals. This result implies that the treated measurements are closer to the true measurement vector than the input measurement vectors. Thus, the data treatment has in this example shown to improve the measurements on the whole.

2.9. Applications using treated data

Discrepancies in overall mass and energy balances are facts for most industrial processes. Few applications can handle deterioration in sensors or the process and still have optimal function. Below follows a list of top applications that benefits from data reconciliation.

- Control
- Instrumentation maintenance
- Overall optimisation and production planning
- Process and sensor diagnostics
- Economy reports
- Environmental reports

The constant striving for higher product quality and lower production costs requires more efficient and accurate control of each variable. Consequently, there is also an increasing demand on sensor accuracy and reliability. Data reconciliation algorithms can lower the maintenance costs by compensating for sensor reading errors and thus keeping performance without physical maintenance until needed to ensure functionality. Consistent data sets after data treatment open up the possibility for advanced control algorithms that have been prevented by deteriorating sensor readings and changing process parameters over time. Overall optimisation is also dependent on consistent data for optimal performance. The impact of data reconciliation is dependent on how sensitive the objective cost function in the optimisation is for faults in data.

Process and sensor diagnostics have much data to mine from the data reconciliation output [Sod00]. Each sensor is continuously monitored and data about deviation from estimated value is reported. Residuals in process models and changes from one sample to another are examples of qualitative data that can be extracted from the data reconciliation. Economy reports for the plant are based upon measurements of produced quantities and raw material in, [Wan01]. The need for consistency in mass and energy balance is obvious to give a reliable economic balance. Some countries have taxation on emissions such as, content of CO₂ or NO_x in flue gas. Connections to costs because of faulty sensor readings are apparent.

3 Results

Three papers and one technical report have been produced to gain insight in when the research hypothesis can be falsified or not. Here the papers are summarised in order to present the most important results. The papers and the report are attached in appendix.

3.1. Paper 1

A heat and power plant need accurate sensor readings to meet efficiency, environmental and safety demands. Unreliable sensor data affect performance and control of the process, causing economical losses. Gross error isolation and data reconciliation is one way for on-line data treatment getting a consistent data set. The system can be put together of different components. In this paper, it was tested how far a very simple approach can reach. With simple it is meant that only already existing instrumentation was used and that some of the fastest components for performing each computational task was chosen for the system.

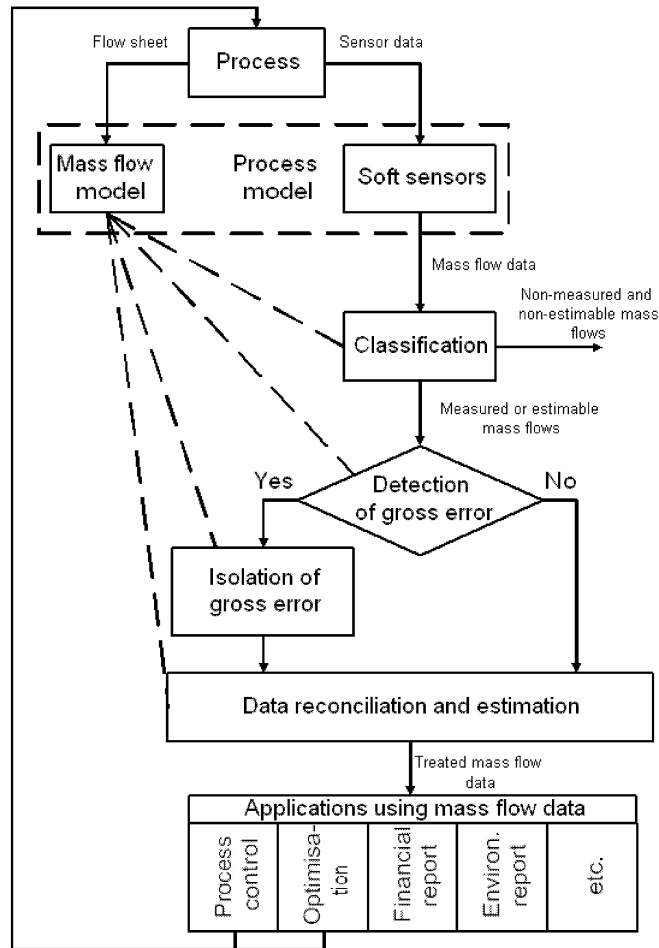


Figure 10. Data treatment from raw data to application.

A linear steady state mass balance model was used in the optimisation algorithm for data reconciliation. A fast hypothesis test (chi-square-test) was used for gross error detection. Single fault was assumed for sequential search in the isolation algorithm. Data reconciliation was performed by analytically solving the weighted linear least squares problem. The following variable estimation was also solved analytically. The system presented in Figure 10 was tested on measurement data from a flue gas channel in a combined heat and power plant in Enköping, Sweden.

The system first calculates physical properties, and then mass flows. The flows are classified into subsets according to redundancy. Redundancy determines to what extent a

sensor can be used in isolation, data reconciliation, and estimation. In the considered flue gas channel there where no hardware redundancy. Limitation to only existing instrumentation led to an exploration of analytical redundancy. It explores not only the hardware redundancy, but also the redundancy that can be found in relations between sensors. These relations are here used in small physical models called soft sensors. Soft sensors provide the data treatment with some of the needed mass flows for being able to perform isolation and data reconciliation. By keeping the soft sensors out of the linear model the core of the system could be kept simple.

	Test				
	1a	2a	2b	3a	3b
Input					
Upper limit	2	2	2	-	-
Lower limit	0.1	0.1	0.1	-	-
Noise	3%	real	real	real	real
Output					
OP	0.302	0.199	0.202	-	-
AVTI	0.041	0.047	0.049	-	-
Selectivity	0.881	0.807	0.805	-	-

Table 5. Summary of performance measures.

Testing of the total system was made in three different test set-ups, where 1 and 2 have different noise, 2a and 2b have different time averages, 3a and 3b also got different time averages. The results are presented in Table 5 in recommended quantities such as overall power (OP), average number of type I error (AVTI) and selectivity. To get data enough for statistical reasons different data sets from steady state operation was put together. The total data set used in the tests contained about 12 000 samples. The same number of samples was used for the simulated tests in test 1a, 2a and 2b. AVTI was kept to between 0.4 – 0.5 to get comparable conditions for the different tests. A gross error between 0.1 to 2 times the average mass flows was added to one of the measured mass flows. The data set with gross error was fed to the gross error isolation and data reconciliation system.

From the results it can be seen that the performance in this case is not affected by averaging of the measurements. A conclusion drawn from the low OP in combination of the high selectivity is that few sensors are detected and isolated while others are not. An analyse of the

individual measurements give that gross errors in the main mass flows are more often detected and isolated than in the low magnitude mass flows. This can be explained by the noise in the main mass flows masking gross errors in the low magnitude mass flows. Maybe scaling of the signals can solve the problem. Other effects not presented is that the unconstrained least squares method produced negative mass flows as well as unreasonable large mass flows, but this can be taken care of with a constrained least squares solver. Until the constrained solver is implemented the estimates are judged as unreliable.

3.2. Paper 2

In this paper the work has been performed together with Björn Widarsson. My part was to, in literature, find the cause-effect relations for fouling between different variables. Björn built the Bayesian network and trained it with a collected data set. In Sweden the use of biomass for power generation in heat and power plants are common, but the experience of power generation with biofuel is limited. The challenge is to combine biomass fuels with other fuels to get a cheap fuel mix and at the same time reduce build-up of fouling. Here we investigated what is affecting fouling in the flue gas channel with focus on the convective superheaters. The different conditions affecting fouling are complex and a tool for decision support is needed. The tool must be able to handle uncertainty in data and be transparent to the user. Transparency is necessary to explain on what basis the decision is taken. Bayesian networks (BN), complies with the wanted qualities and is used for building the tool.

Data for construction of the decision support is collected from the heat and power generation plant in Västerås, Sweden. The decision support is restricted to the convective superheaters that were identified as a problem area. Fouling is a mixture of deposited ash, condensated gasified substances and compounds resulting from reaction between substances on the surface and in the flue gas, that decrease the heat transfer rate in the convective superheaters. Fouling can e.g. be prevented by changing the fuel mix, shorten the time between soot blowing and control boiler temperature. In this work we divided fouling into two categories: hard fouling that cannot be removed by soot blowing and soft fouling that is removed by soot blowing. Calculating the heat transfer coefficient before and after a soot blowing cycle can monitor both hard and soft fouling. We believe that abnormal build-up of fouling can be prevented by strategically choosing fuel mixture and soot blow interval when operating the plant. Thus, shutdown of the plant other than for annual cleaning can be avoided.

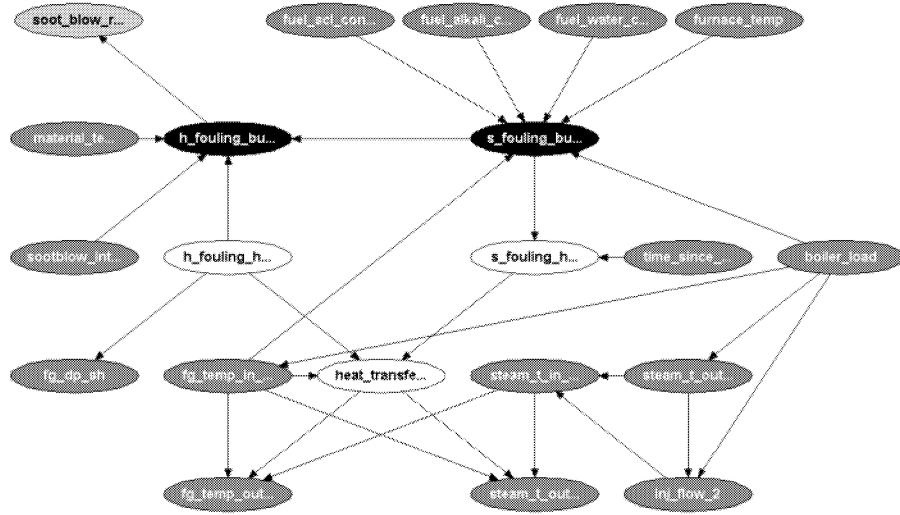


Figure 11. Bayesian network for decision support on soot blowing.

The Bayes theorem [Pea88] together with graph theory is a powerful tool for decision support. The graph is the backbone with the known cause–effect relations. Specific plant data are stored in the condition probability tables. The tables are produced from plant data and a yearlong data set from the distributed control system. The relations between fuel properties and fouling is extracted from technical reports and then put into conditional probability tables. One application of the constructed BN is decision support on soot blowing. Before building this decision support the decision on soot blowing relied on one variable, the steam temperature after the superheater. With a BN many variables can be accounted for in a decision on soot blowing. One of the main features is that preventive actions on fouling can be taken with support of the BN.

The network is built for the superheaters in the flue gas channel and is divided in two parts. The first part is the heat transfer part and the second is the fouling build-up part. The combined BN for decision support on soot blowing is presented in Figure 11. Dark grey nodes represent signals, black nodes are predictions of fouling build-up, recommendation on soot blow is the light grey node (in the top left corner) and white nodes are trained from data. The BN is trained on a yearlong process data set. The rest of the data is used for making test cases to validate the BN. Decision support is given by the soot blow node and it got two states, soot

blow or not soot blow. The decision is taken among others on the maximum hard fouling build up in the end of the season. Test cases are prepared to test different rates of fouling build-up for both soft and hard fouling.

Validation of the results gives that the BN make good prediction of fouling build-up in the normal case when the plant is running at full load. In the low load and middle load region the performance is not so good due to lack of learning data. Important to emphasise is that many variables affecting fouling build-up are not monitored, e.g. fuel properties and particle velocity. Not including all variables affecting fouling reduce the power of the prediction model. However, the BN is quite easily upgraded with more data and additional variables when they become available.

3.3. *Paper 3*

Isolation of gross errors in process plant on-line data is essential for optimal operation. Isolation is normally part of a system including gross error detection, gross error isolation, data reconciliation, and variable estimation. Here we focus on isolation of gross error and dampening in sensors. The aim is to isolate faulty data and correct it on-line. Gross error is defined as a systematic bias of a sensor reading. Dampening is proportional to the amplitude in the part of the sensor reading that is fluctuating. How bias and dampening affect a sine wave is presented in Figure 12.

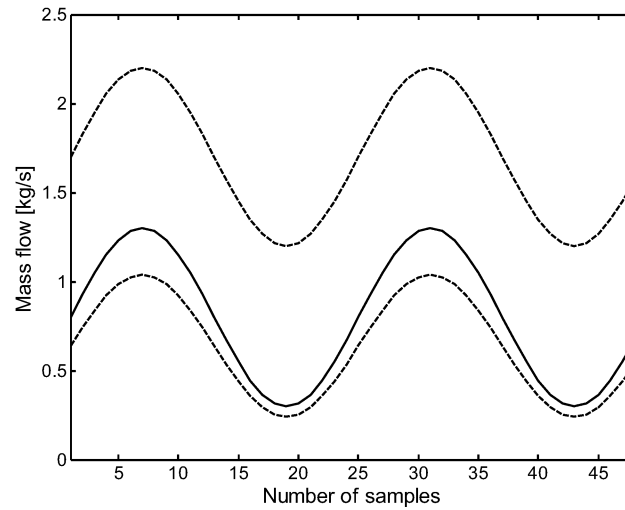


Figure 12. Effect of gross error and dampening for a sine wave. Solid line is true signal, upper dashed line is signal with gross error and lower dashed line is dampened signal.

Two different optimisation approaches are tested for effectiveness in isolation of gross error, δ , and dampening, α . As early detection and isolation as possible is desirable. A time window approach is used to enhance the performance of the algorithms. It is assumed that the time series give extra information to the algorithm and thereby it is less sensitive than an approach using only one sample for isolation. A problem is that real sensor readings contain noise. However, noise cannot be reduced before gross errors and dampening is identified. The noise reduction is performed in a following step called data reconciliation, which is not in the scope of this article. Data reconciliation is not valid for data containing gross errors, and therefore an efficient isolation is essential for a functioning data treatment system.

The test was conducted by adding 1 to 2% relative normal distributed noise to each sensor reading. A time window of 10 samples was created. The same data set was used in all tests.

Gross error or dampening was simulated in one sensor for the algorithm to isolate. The objective is to find the gross error or the dampening that result in a feasible solution of the optimisation problem. The constraints are allowed to give a residual computed from the constraint matrix C , and the estimated variables x . This residual is constrained by the limit, ε .

$$\begin{aligned} \min_{\alpha, \delta, x} & \left(\sum_{i=1}^I \sum_{n=1}^N (x_{i,n} - y_{i,n})^2 \right) \\ \text{s.t.} \quad & -\varepsilon \leq \sum_{n=1}^N c_{m,n} x_{i,n} \leq \varepsilon, \quad m=1, \dots, M, \quad i=1, \dots, I \\ & y_{i,n} = x_{i,n}, \quad n \neq n', \quad i=1, \dots, I \\ & y_{i,n} = \alpha \cdot x_{i,n} + \delta, \quad n = n', \quad i=1, \dots, I \end{aligned} \quad \text{Eq. 3.1}$$

If there is no solution for a given ε , then it is increased in steps until a solution is obtained. Variables where a solution could be found are candidates of containing either gross error or dampening. The solution of the problem gives an estimate of the gross error or the dampening. The second optimisation problem handles noise by introducing a penalty multiplier μ , on the sum of the residual, thus the residual is a penalty in the objective function. In other words the hard limit on the residual in problem Eq. 3.1 is replaced by a soft constraint. The optimisation problem with penalty function can be written as:

$$\begin{aligned} \min_{\alpha, \delta, x} & \left(\sum_{i=1}^I \sum_{n=1}^N (x_{i,n} - y_{i,n})^2 + \mu \sum_{i=1}^I \sum_{m=1}^M \left(\sum_{n=1}^N c_{m,n} x_{i,n} \right)^2 \right) \\ \text{s.t.} \quad & y_{i,n} = x_{i,n}, \quad n \neq n', \quad i=1, \dots, I \\ & y_{i,n} = \alpha \cdot x_{i,n} + \delta, \quad n = n', \quad i=1, \dots, I \end{aligned} \quad \text{Eq. 3.2}$$

The optimisation problem Eq. 3.1 always got a solution. The minimum of the objective function for a given sensor probably contains a gross error or dampening. Solving Eq. 3.2 can

produce the same minimum objective function value for more than one sensor. This is not handled here. An idea is to check the solution for sensors containing unacceptable estimated gross errors or dampening and eliminate them.

For the problems in Eq. 3.1 and Eq. 3.2, the error candidates can be checked afterwards by methods incorporating experience about parameters affecting gross error and dampening. Test results were successful for isolation of absolute gross error. One third, up to half of the sensors was isolated to a unique sensor. Two thirds were not isolated to a unique sensor but to a data set of four sensors up to seven sensors. All of these data sets contained the sensor with the simulated gross error. From the estimated values and gross errors in the solutions it is possible to reduce the number of suspicious sensors in the isolated data set. Dampening was more difficult to isolate. A reduction by 5% of the amplitude was difficult to isolate. Obviously a larger dampening is needed for good performance.

3.4. Report 1

Distributed control systems, DCS, are nowadays widely used in heat and power plants. A historical database combined with the DCS give possibility to use data treatment and diagnostic tools like gross error detection and data reconciliation. Many sensors used for alarm, control and other purposes can also be used by data treatment. One of the restrictions for this report is to use only existing sensors.

Three different model principles are investigated as a base for data treatment. 1) A statistical linear data driven model. 2) A model with mass balance supplemented with soft sensors for mass flow. 3) An energy balance model. Theoretical comparison of the three models resulted in favour for the data treatment based on a mass balance model. The statistical model is preferred for smaller process sections when focus is on fast model development and transparency is not needed. The mass balance model has a base in physical reality and can handle sensor dropouts. The energy balance model got about the same qualities as the mass balance method, however it is more complex and need tuning of parameters to work.

The mass balance model is built for the flue gas channel of the heat and power plant. Other sensors than mass flow sensors were used in equations to calculate mass flows. These blocks for calculation of mass flow are called soft sensors. The core model is used for simulating data and in several operations in the data treatment algorithm. The complete data treatment algorithm is built from standard components such as hypothesis test, classification by projection matrix, sequential isolation and data reconciliation by least squares followed by

data estimation. The purpose of using standard components is to assemble knowledge about the different components for gross error detection, isolation and data reconciliation and how they work together.

The power plant is equipped with more than the standard monitoring instruments, despite that, a number of soft sensors had to be developed to make data reconciliation possible. This point out an important issue: tools for data treatment and diagnosis need to be thought of early, already in the construction phase. This applies to both power plant components and to the complete power plant to ensure observability of critical properties and components.

Results from the simulations show that large errors in mass flow can be detected and to some extent quantified. Mass flow small in magnitude is very hard to detect and isolate when they are connected to larger mass flows. Both the position of the sensor and precision in the process model is important for correct detection and isolation.

4 Concluding remarks

This thesis has tested the hypothesis: quantitative and qualitative methods can deliver a data set in balance from state estimation point of view, despite degradation in process and sensors.

4.1. Conclusions

Measuring mass flow, temperatures and pressures are essential for plant control and monitoring. Heat and power generation plants are suffering from degeneration in sensors and process during operation due to e.g. fouling, erosion and corrosion. These processes benefits from data treatment when it comes to; diagnose purposes, availability of the process, risk management, allocation of maintenance resources, consistent data for statement of accounts and overall control.

Robust and accurate process models are one of the most important module in a data treatment system for reducing gross errors and noise. Process models are usually time-consuming to build, but give both process knowledge and possibilities to simulate process states not practical to test in a real process. Using process model libraries cut engineering effort and time. The accuracy of the model sets the limit for the performance of the tools using the model as a reference. Further investigation of errors in the process or signals need an AI-technique such as BN to produce a detailed diagnose.

Large amounts of measurement data are produced every second in a well-instrumented plant. Measurements can be extracted into information without adding instrumentation hardware by using simulators and data treatment systems. It is important to point out that

necessary instrumentation need to be considered in the construction phase of a plant in order to make full use of process models and diagnose tools.

The data treatment system and its components presented are general for all kinds of flowing processes. The components needed are sensors to measure the state and a process model. The tasks of detection, isolation, and data reconciliation can be solved by different methods and the combinations are numerous.

Hypothesis tests can help to reduce the computation times for a data treatment system by only starting the isolation algorithm when a gross error is detected. Detection tests can be disregarded when the number of sensors is small. The overall power of hypothesis tests are weak for large number of sensors. Hypothesis tests also introduce the problem of false alarms and missed errors, thus the advantages of such test need to be considered before introducing it.

Systematic errors (gross errors) generally make most tools unusable for state estimation and noise reduction. Isolation of gross error is therefore one of the key components in a data treatment system. Using the high accuracy of NLP solutions for isolation is possible due to development of powerful computers and efficient solution algorithms.

Data reconciliation enhances the overall accuracy of the measurements if it is preceded by efficient elimination of gross errors. Degradation in sensors and process affect how to weigh the measurements according to each other in the data reconciliation algorithm. The covariance matrix is often used for this task. This method is fast and easy but not always reliable regarding estimation of covariance from data containing errors. Letting the user set the weights is a solution to the problem, but the updating becomes time-consuming.

Bayesian networks are a promising complement to data reconciliation because of its ability to incorporate both human knowledge in the framework and for the learning of probabilities from data. The network got capability of learning on variables common for the plant operator and thereby opens up the data treatment system for operator interaction.

4.2. Future work

How a system incorporating qualitative and quantitative tools for treating systematic errors in signals and the performance of the tools have been investigated in this thesis.

The possible combinations of tools for data treatment are too many to be investigated systematically by comparative studies. A theoretic study is needed to reveal how different tools can interact to utilise the sensor and expert knowledge in a more efficient manner. Dynamic models can provide useful information for diagnostic purposes and will be

investigated in further work on a thermal plant component. Degradation in process and sensors include both expected and unexpected changes and some degree of adaption is needed. How adaption can be integrated in the total data treatment structure will be studied in a heat and power simulation environment.

5 References

- [Abu02] Abu-el-zeet Z.H., Becerra V.M. and Roberts P.D., Combined bias and outlier identification in dynamic data reconciliation, *Computers and Chemical Engineering*, vol. 26, pp 921-935, **2002**.
- [Ada96] Adams R., *Calculus: a complete course*, Addison Wesley Publishers, **1996**.
- [Ali96] Ali Y. and Narasimhan S., Sensor Network Design for Maximizing Reliability of Bilinear Processes, *American Institute of Chemical Engineers Journal*, vol. 42, pp 2563-2575, **1996**.
- [Alm75] Almassy G. and Sztano T., Checking and correction of measurements on the basis of linear systems model, *Probl. Control Inf. Theory*, vol. 4, p 57, **1975**.
- [Arr03] Arrigada J., *On the Analysis and Fault-Diagnosis Tools for Small-Scale Heat and Power Plants*, PhD Thesis, Lund University, **2003**.
- [Bag00] Bagajewicz M.J., A brief review of recent developments in data reconciliation and gross error detection/estimation, *Latin American Applied Research*, vol. 30, pp 335-342, **2000**.
- [Bag03] Bagajewicz M.J. and Cabrera E., Data Reconciliation in Gas Pipeline Systems, *Industrial and Engineering Chemical Research*, vol. 42, pp 5596-5606, **2003**.
- [Bag97] Bagajewicz M., Optimal Sensor Location in Process Plants, sensor placement, *American Institute of Chemical Engineers Journal*, vol. 43, p 2300, **1997**.
- [Che98] Chen J. and Romagnoli J., A strategy for simultaneous dynamic data reconciliation and outlier detection, *Computers and Chemical Engineering*, vol. 22, pp 559-562, **1998**.
- [Cro83] Crowe C.M., Campos Y.A.G. and Hrymak A., "Reconciliation of process flow rates by matrix projection. Part I: Linear case, *American Institute of Chemical Engineers Journal*, vol. 29, pp 881-888, **1983**.
- [Cro86] Crowe C.M., "Reconciliation of Process Flow Rates by Matrix Projection, Part II: The Nonlinear Case, *American Institute of Chemical Engineers Journal*, vol. 32, pp 616-623, **1986**.
- [Cro96] Crowe C.M., Data reconciliation – progress and challenges, *Journal of Process Control*, vol. 6, pp 89-98, **1996**.
- [Deo74] Deo N., *Graph Theory with Applications to Engineering and Computer Science*, Englewood Cliffs, Prentice-Hall, **1974**.
- [Doy72] Doyle W.H., Instrument-connected Losses in the CPI, *Instrumentation Technology*, vol. 19, pp 38-40, **1972**.

- [Dun96] Dunia R., Qin S.J., Edgar T.F. and McAvoy T.J., Identification of Faulty Sensors Using Principal Component Analysis, *American Institute of Chemical Engineers Journal*, vol. 42, pp 2797-2812, **1996**.
- [Fre00] Frennelius A., *Sannolikhetslära och statistisk inferens för tekniska utbildningar*, 5th edition, Västerås, Sweden, Västerås Statistikutbildning, **2000**, (in Swedish).
- [Gat02] Gatzke E.P. and Doyle F.J. III, "Use of multiple models and qualitative knowledge for on-line moving horizon disturbance estimation and fault diagnosis, *Journal of Process Control*, vol. 12, pp 339-352, **2002**.
- [Gri92] Grierson D.E., Rzevski G. and Adey R.A., Applications of Artificial Intelligence in Engineering VII, *Computational Mechanics Publications*, Southampton, Boston, USA, **1992**.
- [Hog78] Hogg R.V. and Craig A.T., *Introduction to Mathematical Statistics*, 4th edition, New York, Macmillan, **1978**.
- [Hoo04] Hood E., Lessons Learned?, *Environmental Health Perspectives*, vol. 112, pp A352-A358, **2004**.
- [Jia99] Jiang Q., Sanchez M. and Bagajewicz M.J., On the Performance of Principal Component Analysis on Multiple Gross Error Detection, *Industrial and Engineering Chemical Research*, vol. 38, pp 2005-2012, **1999**.
- [Jor87] Joris P. and Kalitventzeff B., Process measurements analysis and validation, Proceedings of CEF'87: *Use Computational Chemical Engineering*, Italy, pp 41-46, **1987**.
- [Kal60] Kalman R.E., A New Approach to Linear Filtering and Prediction Problems, *Transactions ASME Journal of Basic Engineering*, vol. 82D, pp 35-45, **1960**.
- [Kes97] Kesavan P. and Lee J.H., Diagnostic tools for multivariable model-based control systems, *Industrial and Engineering Chemistry Research*, vol. 36, pp 2725-2738, **1997**.
- [Kle98] Kletz T.A., "What went wrong – Case histories of process plant disasters, Gulf Publishing, **1998**.
- [Kne80] Knepper J.C. and Gorman J.W., Statistical Analysis of Constrained Data Sets, *American Institute of Chemical Engineers Journal*, vol. 26, pp 260-264, **1980**.
- [Kue61] Kuehn D.R. and Davidson H., Computer control II mathematics of control. *Chemical Engineering Progress*, vol. 57, pp 44-47, **1961**.
- [Lef92] Lefebvre B. and Santyr G., *Chemical Recovery in the Alkaline Pulping Processes*, Atlanta, TAPPI, 79, **1992**.
- [Lie91] Liebman M.J., *Reconciliation of Process Measurements using Statistical and Nonlinear Programming Techniques*, PhD Thesis, The University of Texas and Austin, **1991**.

- [Lie92] Liebman M.J., Edgar T.F. and Lasdon L.S., Efficient Data Reconciliation and Estimation for Dynamic Processes using Nonlinear Programming Techniques, *Computers and Chemical Engineering*, vol. 16, pp 963-986, **1992**.
- [Mad92] Madron F., *Process Plant Performance. Measurement and Data Processing for Optimisation and Retrofits*, Ellis Horwood, Chichester, England, **1992**.
- [Mah76] Mah R.S.H., Stanley G.M. and Downing D.M., Reconciliation and rectification of process flow and inventory data, *Industrial Engineering Chemistry Process Design Development*, vol. 15, p 175-183, **1976**.
- [Mah82] Mah R.S.H. and Tamhane A.C., Detection of Gross Errors in Process Data, *American Institute of Chemical Engineers Journal*, vol. 28, pp 828-830, **1982**.
- [Mak95] Makni S., Hodouin D. and Bazin C., A Recursive Node Imbalance Method Incorporating a Model of Flowrate Dynamics for On-line Material Balance of Complex Flowsheets, *Minerals Engineering*, vol. 8, pp 753-766, **1995**.
- [Mar01] Marshall J., L_1 -norm pre-analysis measures for geodetic networks, *Journal of Geodesy*, vol. 76, pp334-344, **2002**.
- [Mat02] *Statistics Toolbox, Matlab user's guide version 4*, Mathworks, **2002**.
- [McB95] McBrayer K. and Edgar T.F., Bias detection and estimation in dynamic data reconciliation, *Journal of Process Control*, vol. 5, pp 285-289, **1995**.
- [Mor99] Morad K., Svrcek W.Y. and McKay I., A robust direct approach for calculating measurement error covariance matrix, *Computers and Chemical Engineering*, vol. 23, pp 2725-2738, **1999**.
- [Nar00] Narasimhan S., Jordache C., *Data Reconciliation & Gross Error Detection, An Intelligent Use of Process Data*, Gulf Publishing Company, Houston, USA, **2000**.
- [Nar87] Narasimhan S. and Mah R.S.H., Generalized Likelihood Ratio Method for Gross Error Identification, *American Institute of Chemical Engineers Journal*, vol. 33, pp 1514-1521, **1987**.
- [Nog72] Nogita S., Statistical test and adjustment of process data, *Industrial Engineering Chemistry Process Design Development*, vol. 11, pp 197-200, **1972**.
- [Ols94] Olsen O.A., *Instrumenteringsteknikk – med vekt på måleteknikk*, Tapir, Trondheim, **1994**, (in Norwegian).
- [Pea88] J. Pearl, *Probabilistic Reasoning in Intelligent Systems – Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, California, USA, **1988**.

- [Rao99] Rao R., Bhushan M. and Rengaswamy R., Locating Sensors in Complex Chemical Plants Based on Fault Diagnostic Observability Criteria, *American Institute of Chemical Engineers Journal*, vol. 45, pp 310-322, **1999**.
- [Rei63] Reilly P.M. and Carpani R.E., Application of Statistical Theory of Adjustments on Material Balances, presented at the *13th Canadian Chemical Engineering Conference*, Ottawa, Canada, **1963**.
- [Rip65] Ripps D.L., Adjustment of Experimental Data, *Chemical Engineering Progress Symposium*, Series 61, pp 8-13, **1965**.
- [Rol02] Rollins D.K., Devanathan S. and Bascuñana M.V.B., Measurement bias detection in linear dynamic systems, *Computers and Chemical Engineering*, vol. 26, pp 1201-1211, **2002**.
- [Rol92] Rollins D.K. and Davis J.F., Unbiased Estimation of Gross Errors in Process Measurements, *American Institute of Chemical Engineers Journal*, vol. 38, pp 563-572, **1992**.
- [Rol96] Rollins D.K. Cheng Y. and Devanathan S., Intelligent Selection of Hypothesis Test to Enhance Gross Error Detection, *Computers and Chemical Engineering*, vol. 20, pp 517-530, **1996**.
- [Rom00] Romagnoli J.A., Sanchez C.S., *Process systems engineering volume 2, Data Processing and Reconciliation for Chemical Process Operations*, Academic Press, San Diego, USA, **2000**.
- [Rom81] Romagnoli J. and Stephanopoulos G., Rectification of process measurement data in the presence of gross errors, *Chemical Engineering Science*, vol. 36, pp 1849-1863, **1981**.
- [Rom83] Romagnoli J., On data reconciliation constraints processing and treatment of bias, *Chemical Engineering Science*, vol. 38, pp 1107-1117, **1983**.
- [Rus00] Russell E.L., Chiang L.H. and Braatz R.D., *Data-driven techniques for fault detection and diagnosis in chemical processes*, Springer, London, **2000**.
- [Rus99] Russo L.P. and Yang R.E., Moving Horizon State Estimation Applied to an Industrial Polymerization Process, *Proceedings of the American Control Conference*, June **1991**.
- [San92] Sánchez M., Bandoni A. and Romagnoli J., PLADAT-A package for process variable classification and plant data reconciliation. *Computers and Chemical Engineering*, S16, pp 499-506, **1992**.
- [San96] Sánchez M. and Romagnoli J., Use of orthogonal transformations in data classification and reconciliation, *Computers and Chemical Engineering*, vol. 20, pp 483-493, **1996**.

- [Ser86] Serth R. and Heenan W., Gross error detection and data reconciliation in steam metering systems, *American Institute of Chemical Engineers Journal*, vol. 32, pp 733-742, **1986**.
- [Sim91] Simpson D.E., Voller V.R. and Everett M.G., An Efficient Algorithm for Mineral Processing Data Adjustment, *International Journal of Mineral Processing*, vol. 31, pp 73-96, **1991**.
- [Sod00] Soderstrom T.A., Thomas F.E., Louis P.R. and Yang R.E., Industrial Application of a Large-Scale Dynamic Data Reconciliation Strategy, *Industrial and Engineering Chemical Research*, vol. 39, pp 1683-1693, **2000**.
- [Sod01] Soderstrom T.A., Himmelblau D.M. and Edgar T.F., A mixed optimization approach for simultaneous data reconciliation and identification of measurement bias, *Control Engineering Practice*, vol. 9, pp 869-876, **2001**.
- [Sun03] Sunde S., Berg Ø., Dahlberg L. and Fridqvist N., Data Reconciliation in the Steam-Turbine Cycle of a Boiling Water Reactor, *Nuclear Technology*, vol. 143, pp 103-123, **2003**.
- [Swa89], Swartz C.L.E., Data Reconciliation for Generalized Flowsheet Applications, *American Chemical Society National Meeting*, Dallas, Tex, **1989**.
- [Tam88a] Tamhane A.C., A Bayesian Approach To Gross Error Detection in Chemical Process Data, Part I : Model Development, *Chemometrics and Intelligent Laboratory Systems*, vol. 4, pp 33-45, **1988**.
- [Tam88b] Tamhane A.C., A Bayesian Approach To Gross Error Detection in Chemical Process Data, Part II : Simulation Results, *Chemometrics and Intelligent Laboratory Systems*, vol. 4, pp 131-146, **1988**.
- [Ter93] Terry P.A. and Himmelblau D.M., Data Rectification and Gross Error Detection in a Steady-State Process via Artificial Networks, *Industrial and Chemical Engineering*, vol. 32, pp 3020-3028, **1993**.
- [Ton94] Tong H. and Crowe C.M., The application of principal component analysis to tests for gross error in data reconciliation, *The 5th International Symposium on Process Systems Engineering*, Kyongju, Korea, **1994**.
- [Vac69] Václavek V., Studies on System Engineering III, Optimal Choice of the Balance Measurements in Complicated Chemical Engineering Systems, *Chemical Engineering Science*, vol. 24, pp 947-955, **1969**.
- [Vac72] Václavek, V. Testing of Over- and Underdetermination of Problem Formulation for Simulation Calculations of a Complex Chemical Process, *Chemical Engineering Science*, vol. 27, pp 1669-1676, **1972**.

- [Vac76] Václavek V. and Loucka M., Selection of Measurements Necessary to Achieve Multicomponent Mass Balance in Chemical Plant, *Chemical Engineering Science*, vol. 31, pp 1199-1205, **1976**.
- [Vat73] Vatslavek V., Kubichek M. and Louchka M., "Calculation of material balances in chemical engineering systems with allowance for errors in measurement classification of stream parameters. Theoretical foundations of chemical engineering, a cover-to-cover translation of *Teoreticheske osnovy chimiceskoj technologii*, New york, N.Y. Consultants bureau, Plenum Publishing Corporation, pp 256-260, **1976**.
- [Ven03a] Venkatasubramanian V., Rengaswamy R., Yin K. and Kavuri S.N., A review of process fault detection and diagnosis, Part I: Quantitative model-based methods, *Computers and Chemical Engineering*, vol. 27, pp 293-311, **2003**.
- [Ven03b] Venkatasubramanian V., Rengaswamy R., Yin K. and Kavuri S.N., A review of process fault detection and diagnosis, Part II: Qualitative models and search strategies, *Computers and Chemical Engineering*, vol. 27, pp 313-326, **2003**.
- [Ven03c] Venkatasubramanian V., Rengaswamy R., Yin K. and Kavuri S.N., A review of process fault detection and diagnosis, Part III: Process history based methods, *Computers and Chemical Engineering*, vol. 27, pp 327-346, **2003**.
- [Wan01] Wang Y., Rong G. and Wang S., Linear Dynamic Data Reconciliation: Refinery Application, *6th IFAC Symposium on Dynamics and Control of Process Systems*, Korea, pp 650-655, **2001**.
- [Wat89] Watson I.A. and Oakes F., *Human reliability factors in technology management*, Measurement and Technology, 1989, vol. 136, issue 1, pp24-32. **1989**.
- [Wei01] Weidl G., *Root Cause Analysis and Decision Support on Process Operation*, Dissertation, Mälardalen University Press, Västerås, Sweden, **2001**.
- [Wei04] Weisstein E.W., *Degree of Freedom*, From MathWorld - A Wolfram Web Resource. <http://mathworld.wolfram.com/DegreeofFreedom.html>
- [Wil74] Wilsky A. S. and Jones L., A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems, *IEEE Transactions on Automatic Control*, vol. 21, pp 108-112, **1974**.
- [Wil99] Williams P., *Model building in mathematical programming*, 4th edition, John Wiley and Sons, **1999**.
- [Win95] Winston W.L., *Introduction to Mathematical Programming – Applications and Algorithms*, Belmont, California, Duxbury Press, **1995**.