



University
of Glasgow

Xia, Chunming (2003) *Control loop measurement based isolation of faults and disturbances in process plants.*

PhD thesis

<http://theses.gla.ac.uk/4080/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

**CONTROL LOOP MEASUREMENT BASED
ISOLATION OF FAULTS AND DISTURBANCES
IN PROCESS PLANTS**

CHUNMING XIA

**A THESIS SUBMITTED TO
THE DEPARTMENT OF MECHANICAL ENGINEERING OF
THE FACULTY OF ENGINEERING OF
THE UNIVERSITY OF GLASGOW FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY**

©CHUNMING XIA, MARCH 2003, GLASGOW, SCOTLAND

Acknowledgements

First of all I would like to express my deep gratitude to my supervisor, Dr. John Howell, for his continuous help and guidance.

I would like to give my thanks to Mrs. Nina F. Thornhill from UCL, for her valuable advice, help and supply of data; to John Cox and Michael Paulonis from *Eastman Chemical Company, Kingsport, TN, USA*, for the supply of real plant data; and to Dr. Jun Chen, Dr. Wen-Hua Chen, Dr. Zhihong Huang and Dr. Yongji Wang, for their friendship and encouragement.

Thanks also go to all the other people who have helped me during my stay at the University of Glasgow.

Special thanks go to my sisters, brothers and their families, for their love and supports.

I must thank my wife, Yihong, and my son, Jiasheng, for their love.

To my wife and my son

In memory of my parents

Contents

List of Tables	v
List of Figures	vi
Abstract	ix
Chapter 1: Introduction	1
1.1 Background.....	1
1.2 Aims and Objectives.....	5
1.3 Testing the Techniques that Meet These Aims and Objectives.....	7
1.4 Outline of the Thesis.....	8
1.5 Originality.....	10
Chapter 2: Some Background to Process Monitoring and Fault Detection & Isolation	11
2.1 Overview.....	11
2.2 Control Loop Performance Assessment	16
2.2.1 MVC Benchmark Based Approaches	16
2.2.2 Modifications, Extensions and Other Non-MVC Based Approaches .	22
2.2.3 Fault Detection Based on Performance Index	23
2.2.4 Criticisms	23
2.3 Oscillation Detection & Diagnosis	25
2.3.1 Common Causes of Oscillations.....	25

2.3.2	Single-Loop Based Approaches.....	26
2.3.3	Analysis of Plant-Wide Oscillations.....	32
Chapter 3: Loop Status Monitoring		35
3.1	Overview.....	35
3.2	Fundamentals	39
3.2.1	The Model.....	39
3.2.2	The Basis	41
3.2.3	Estimation of η_y , η_u , R_n and Standard Deviation of R_n	50
3.3	Loop Status Monitoring	53
3.3.1	The Qualitative Loop Status Statistic	53
3.3.2	The Quantitative Loop Status Statistic	59
3.4	Examples.....	61
3.5	Summary.....	64
Chapter 4: Fault Localisation in a Number of Interacting Control Loops		66
4.1	OLPI Derivations & Fault Localisation.....	67
4.2	Application to Real Data From an Eastman Chemical Company Plant	69
4.3	Case Studies with the Tennessee Eastman Process Benchmark.....	72
4.3.1	Case 1: Step Change Disturbance to Loop 5 (B Composition)	74
4.3.2	Case 2: A-Feed Flow Sensor Drift.....	76
4.3.3	Case 3: Reaction Kinetic Slow Drift.....	77
4.4	Summary.....	79

Chapter 5: The Case When Two Loops Have Similar Significant

OLPI Values	80
5.1 Tightly Coupled Control Loops	80
5.2 Commissioning-Stage Poor Performance (CSPP)	83
5.2.1 Poor Design Vs. a Commissioning-Stage Fault	83
5.2.2 Test of Control Efforts	84
5.2.3 Test of Uniqueness of Dominant Spectral Components	86
5.2.4 Test of Physical Properties	86
5.3 Case Studies	88
5.3.1 Case1: Tightly Coupled Through Recycling	88
5.3.2 Case 2: Valve Problem in LC1	91
5.4 Conclusions	92

Chapter 6: Detecting and Isolating Multiple Plant-Wide

Oscillations	93
6.1 Introduction	94
6.2 Methods	96
6.2.1 Time Domain ICA	96
6.2.2 Basic Spectral ICA	99
6.2.3 Some Practicalities	105
6.2.4 Dominance of ICs and Isolation of Sources	106
6.2.5 Multi-Range Spectral ICA	107
6.3 Examples	111
6.3.1 Example1: Simulated Plant Data	111
6.3.2 Example2: Case Study on the Industrial Data Set	117

6.4 Conclusions.....	127
Chapter 7: Conclusions and Recommendations.....	129
7.1 Conclusions.....	129
7.2 Future Work.....	132
References.....	133
Appendix A: Proofs for the Basis in Chapter 3	150
A.1 Proof of Theorem 3-1.....	150
A.2 General Deterministic Oscillatory Trends (PI).....	153
A.3 Oscillatory Trend with Frequency ω (PID)	155
Appendix B: Analysis Results of Real Data Set From an Eastman Chemical Company Plant	157
Appendix C: The Model of 2-Cascaded CSTRs	166
C.1 Up CSTR Tank Model.....	166
C.2 Down CSTR Tank Model.....	169
C.3 Operating Data for Two CSTRs	171
Appendix D: The Kurtosis of Some Typical Power Spectra.....	176

List of Tables

Table 2-1: Recommended CLPA parameters	20
Table 3-1: Some possible causes for loop statuses	38
Table 3-2: Sensitivity analysis of Q	45
Table 3-3: Loop Status criteria	53
Table 3-4: Quantitative values of Loop Status	59
Table 3-5: Rules for converting the quantitative statistic into a qualitative version	60
Table 4-1: Description of 9 master loops in an Eastman Chemical Company plant	69
Table 4-2: <i>OLPIs</i> & Loop Status for an Eastman Chemical Company plant real data set	71
Table 4-3: The nine master loops of the TE process benchmark.....	73
Table 5-1: Criteria of physical test	87
Table 5-2: <i>OLPI</i> Values for TC1 &TC2.....	88
Table 5-3: <i>OLPI</i> values for LC1 & LC2 with valve problem in LC1 loop	91
Table 6-1: Values of significance indexes, dominance ratios and estimated periods for example 1	114
Table 6-2: Separation of the multi-frequency ranges and their associated $\alpha\%$ parameters.....	120
Table 6-3: Dominance ratio (DR) and period results of example 2.....	121
Table C-1: Steady state values & parameter values for the up CSTR tank.....	168
Table C-2: Parameters for up CSTR controllers.....	168
Table C-3: Time delays for up CSTR control processes	169
Table C-4: Steady state values & parameter values for the down CSTR tank	170
Table C-5: Parameters for down CSTR controllers.....	170

List of Figures

Figure 2-1: Schematic diagram of SISO process under feedback control.....	17
Figure 2-2: CLPA index plot with respect to prediction horizon d	21
Figure 2-3: Ideal $y(t)$ and $u(t)$ signals in the presence of stiction.....	29
Figure 3-1: Loop with deterministic disturbance & white noise excitation	40
Figure 3-2: Filtered white noise frequency response for a typical PI control loop .	41
Figure 3-3: η_u , R relationships with ω when $\eta_y=0.5$	46
Figure 3-4: Plots of R_n versus $T_i\omega$ for PI controllers	47
Figure 3-5: Plots of R_n versus $T_i\omega$ for PID controllers	50
Figure 3-6: Results of pdf($\hat{R}_n(k)$)	55
Figure 3-7: Some simple examples of loop statuses.....	57
Figure 3-8: Flowchart of loop status assessment.....	58
Figure 3-9: Examples of filtered LL statistics and icon statuses.....	62
Figure 3-10: Comparison between a low- f oscillation and a non-stationary disturbance.....	63
Figure 4-1: Time trends of 9 master loops in an Eastman Chemical Company plant	70
Figure 4-2: Flow diagram and basic control of the T-E process benchmark.....	73
Figure 4-3: Step change disturbance to B composition: Loop 5 data.....	75
Figure 4-4: Step change disturbance to B composition: $OLPI$ & LL	75
Figure 4-5: A-feed flow sensor drift: Loop 1 data.....	76
Figure 4-6: A-feed flow sensor drift: $OLPI$ & LL	76
Figure 4-7: Reaction kinetic slow drift: Loop 3 data.....	77
Figure 4-8: Reaction kinetic slow drift: $OLPI$	78

Figure 5-1: Plant schematic of two cascaded CSTRs	82
Figure 5-2: The power spectra of TC1 & TC2 controller outputs	89
Figure 5-3: Power trends and cross correlation	90
Figure 6-1: The simulation plant schematic	112
Figure 6-2: Normalised loop trends and normalised power spectra	113
Figure 6-3: Independent components (ICs) and principal components (PCs) of example 1	114
Figure 6-4: Significance index plot for IC1 and IC2 of example 1	115
Figure 6-5: Normalised PVs and power spectra	118
Figure 6-6: Normalised OPs and power spectra	119
Figure 6-7: Independent components of example 2	121
Figure 6-8: Significance index plot for low frequency IC2, IC4 and IC5	122
Figure 6-9: Other significance index plots for ICs in different frequency ranges .	122
Figure B-1: Results for loop PC1	157
Figure B-2: Results for loop LC1	158
Figure B-3: Results for loop FC4	159
Figure B-4: Results for loop TC1	160
Figure B-5: Results for loop PC2	161
Figure B-6: Results for loop LC3	162
Figure B-7: Results for loop LC2	163
Figure B-8: Results for loop TC2	164
Figure B-9: Results for loop FC7	165
Figure C-1: Simulated up stream CSTR process	166
Figure C-2: Simulated down stream CSTR process	169
Figure C-3: Controlled variables & controller outputs (up CSTR)	171
Figure C-4: Other measurements (up CSTR)	172

Figure C-5: Controlled variables & controller outputs (down CSTR)	172
Figure C-6: Other measurements (down CSTR)	173
Figure C-7: Controlled variables & controller outputs (up CSTR, linkage)	173
Figure C-8: Other measurements (up CSTR, linkage)	174
Figure C-9: Controlled variables & controller outputs (down CSTR, linkage).....	174
Figure C-10: Other measurements (down CSTR, linkage).....	175
Figure D-1: Kurtosis of the power spectrum of a deterministic trend with 2 sine waves	181

Abstract

Global commercial pressures have resulted in an increased demand on process control engineers to detect and diagnose control loop problems in chemical process plants quickly as these can affect plant performance. A large petrochemical plant may have a thousand or more control loops and indicators, so a key requirement of any industrial control engineer is for an automated means to detect and isolate the root causes so that maintenance effort can be directed efficiently. Detection and diagnosis of plant-wide oscillations are of particular importance, because the propagation of an oscillation throughout the plant can have an impact on product quality and running costs.

Although there has been considerable commercial and academic interest in methods for analysing the performance of control systems, they are usually designed to detect, and not to isolate or diagnose, faults or other root causes of poor performance. The most successful methods compare variations in the controlled variable to those that would be expected if the loop is under minimum variance control (MVC). They assume that the plant is in a steady state, in which the only variation in the controlled variable of a closed loop is as a result of stochastic noise. Automated implementations can often trigger false alarms, due to the existence of transient and/or external deterministic disturbances in closed control loops. In particular they are susceptible to whole plant oscillations.

This thesis focuses on the development of data-driven automated techniques to enhance these performance assessment methods. These techniques include process control loop status monitoring, fault localisation in a number of interacting control loops and the detection and isolation of multiple oscillations in a multi-loop situation. Not only do they make use of controlled variables, but they also make use of controller outputs, indicator readings, set-points, and controller settings.

The idea behind *loop status* is that knowledge of the current behaviour of a loop is important when assessing MVC-based performance, because of the assumptions that are made in the assessment. Current behaviour is defined in terms of the kind of deterministic trend that is present in the loop at the time of assessment. When the status is other than steady, MVC-based approaches are inappropriate. Either the assessment must be delayed until steady conditions are attained or other methods must be applied. When the status is other than steady, knowledge of current behaviour can help identify the possible cause. One way of doing this is to derive another statistic, the overall loop performance index (OLPI), from loop status. The thesis describes a novel fault localisation technique, which analyses this statistic to find the source of a plant-wide disturbance, when a number of interacting control loops are perturbed by a single dominant disturbance/fault. Although the technique can isolate a single dominant oscillation, it is not able to isolate the sources of multiple, dominant oscillations. To do this, a novel technique is proposed that is based on the application of spectral independent component analysis (ICA).

Both simulated and real plant data are used to test the proposed approaches. The results show that loop statuses can provide current trend information about PI/PID control loop performance, and help to narrow down the possible causes. OLPI-based fault localisation techniques are successful when analysing a plant with interacting loops that is perturbed by a dominant, deterministic disturbance. A case study on a tightly coupled two-CSTR simulated system study shows that it is capable of distinguishing between a poor design problem and a commissioning-stage fault.

Spectral independent component analysis (spectral ICA) is based on the analysis of spectra derived via a discrete Fourier transform from time domain process data. The analysis is able to extract dominant spectrum-like independent components each of which has a narrow-band peak that captures the behaviour of one of the oscillation sources. It is shown that the extraction of independent components with single spectral peaks can be guaranteed by an ICA algorithm that maximises the kurtosis of the independent components (ICs). This is a significant advantage over spectral principle component analysis (PCA), because multiple spectral peaks could be present in the extracted principle components (PCs), and the interpretation of detection and isolation of oscillation disturbances based on spectral PCs is not straightforward. The novel spectral ICA method is applied to a simulated data set and to real plant data obtained from an industrial chemical plant. Results demonstrate its ability to detect and isolate multiple dominant oscillations in different frequency ranges.

Chapter 1

Introduction

1.1 Background

It is important for process control engineers to detect and diagnose control loop problems in chemical process plants quickly (Paulonis & Cox 2003) as these can affect plant performance. For instance they can induce oscillations that increase variability and can prevent a plant from operating close to optimal constraints. They can also camouflage other behaviour that may need attention such as upsets due to external disturbances. A large petrochemical plant may have a 1000 or more control loops and indicators, so a key requirement of an industrial control engineer is for an automated means to detect and isolate the root causes so that maintenance effort can be directed efficiently.

Over the past decade there has been considerable commercial and academic interest in methods for analysing the performance of a controller. The development of minimum variance control (MVC) benchmark based, off-line, closed loop, performance assessment techniques (Harris & Seppala 2001; Huang & Shah 1999) are now so well established that various vendors are offering commercial analysis products based on them. For example, Honeywell offer the *Loop ScoutTM* software package, and Matrikon offer *ProcessDoctor*. These performance assessment techniques are usually designed to detect, and not to isolate or diagnose, faults or other root causes of poor performance. As highlighted by Harris and Seppala

(2001), a comprehensive approach for controller performance monitoring should also include the development of methods for diagnosing the underlying causes for changes in the performance of a control system. Another important limitation is that the assessment assumes that the plant is in a steady state, in which the only variation in the controlled variable of a closed loop is as a result of stochastic noise. This condition is not satisfied in many plant situations. Thus the automated performance monitoring toolkits can often trigger false alarms. Such false alarms will most likely occur when either a loop is operating transiently, or it is operating in a steady state that exhibits a dominant deterministic trend like an oscillation. The latter is quite common in many chemical plants where a limit cyclic oscillation can propagate to other process loops causing them to false alarm. Yet another weakness of the MVC-based performance index is that it is what is known as a grey indicator for fault detection because sometimes it is unable to detect a fault. For example, a PI loop can effectively compensate for a slow sensor drift, because this effect is unlikely to be visible in the controlled variable to any extent.

Thus the development of an automated tool that can reduce the false alarm rate, enhance the information available about the current function of a loop and isolate the problem loop in an interacting plant that contains a number of perturbed loops, is important.

Some attempts have been made to enhance performance assessment with methods to detect and diagnose oscillations in control loops. Most of these methods focus on SISO control loops, although the detection and clustering of plant-wide oscillations that contain more than one fundamental frequency (i.e. multiple

oscillations) has also been studied (Thornhill *et al.* 2002a; Thornhill *et al.* 2003). Little has been done to isolate the sources of such multiple oscillations.

Many practitioners, in the area of chemical process monitoring, have recently discussed the need for automated tools that can provide effective and practical techniques to meet the need of large-scale process monitoring and diagnosis.

- Kozub (1997) of the Shell Development Company has compared experiences with direct time series based approaches to minimum variance based techniques. Minimum variance based techniques focus on the analysis of time series that pertain to one or more statistics. He observed that although the use of a single number statistic can hardly be as effective as analysis tools that offer far more detail, the latter are very resource intensive whilst a group of single number statistics would be useful to provide a first pass indication that a problem exists. Kozub sees the potential for both detecting and determining the extent of under-damped, or cyclical, response trend characteristics and views the automation of controller performance monitoring and diagnosis as an important new challenge.
- Paulonis and Cox (2003) of the Eastman Chemical Company recently wrote that there is significant interest in improving disturbance diagnosis by identifying sets of loops that appear to share a common disturbance and even identifying a loop that may be the root cause of the distributed disturbance.

- Fedenczuk (BP Grangemouth) recently emailed the following: “There continues to be lack of something simple, which gives an indication whether a loop's performance is healthy. If the loop fails the initial health performance check then a secondary detailed analysis is required to identify the cause: instrument, tuning, valve position, process disturbance etc.”.
- Desborough *et al.* (2001) at Honeywell have claimed that “process control engineers don't have the tools to allow them to continuously monitor all of their loops. After 10 years of studying control loops at hundreds of sites, we know process control has a long way to go, despite significant technical advances in performance monitoring”.

1.2 Aims and Objectives

This thesis describes control loop measurement based approaches for (1) process control loop status monitoring, (2) fault localisation in a number of interacting control loops, (3) the analysis of poor performance caused by tightly coupled loops due to poor plant design, and (4) the detection and isolation of multiple oscillations in a multi-loop situation. These approaches are mainly data-driven, i.e. they are based on available control loop measurements and controller settings. No detailed process model information is needed. Measurements may include:

- controlled variables;
- controller outputs;
- sensor readings from indicators;
- set-points.

Their combined aims are to:

- indicate what kind of deterministic trend is exhibiting in a loop, and, if abnormal, what are some possible causes;
- isolate the root-cause loop in a number of interacting loops, which contain abnormal deterministic trends;
- detect and isolate multiple plant-wide oscillations, which are the most common causes for the severe deterioration of loop performance;
- accommodate special cases of poor performance, which might be caused by tightly-coupled loops due to poor process and/or control design, and might arise at commissioning when normal operation history is not available.

The various approaches are listed below.

- (1) The current function of a loop is characterised in terms of its *loop status* indicating what kind of deterministic trend is present in the loop.
- (2) Loop status is used to form the basis for the development of a fault localisation technique, which can find the source of a plant-wide disturbance, when a number of interacting control loops are perturbed by a single dominant disturbance/fault.
- (3) Tests on net energy/mass trends are developed for the analysis of the special cases described above.
- (4) Spectral independent component analysis (ICA) is developed to detect and isolate multiple oscillations.

By focusing on 'current' loop status rather than on an audit of loop performance, loop status monitoring can be viewed as a step towards Kozub's vision (1997).

1.3 Testing the Techniques that Meet These Aims and Objectives

Fundamental to this thesis is the need to test on data because the techniques are data-driven. Data was obtained from four sources.

- (1) Crude and simple simulations were used to generate data to examine basic properties.
- (2) A simulation of a tightly-coupled system, containing two cascaded CSTRs, was constructed (see Appendix C), to provide data to examine the effects of either poor process design or component faults that cause the data to have similar properties.
- (3) The Tennessee Eastman (T-E) process benchmark provided realistic data with different fault conditions.
- (4) Real data was obtained from an Eastman Chemical Plant. This contained multiple plant-wide oscillations. The most dominant oscillation had already been diagnosed by plant engineers, whilst other sources of oscillations were yet to be isolated.

1.4 Outline of the Thesis

Chapter 2 first briefly reviews some general methods for process monitoring and fault detection & isolation (FDI), and then focuses on some specific methods that have had direct influences on the development of the approaches proposed.

A new concept, that of *Loop Status* monitoring, is then described in Chapter 3. The idea is that the status would indicate what kind of deterministic trend is present in a PI/PID loop and, if abnormal, what is the possible cause. The chapter emphasises (1) the categorisation of a deterministic trend as one of seven statuses; (2) the scientific basis of the statistics needed to perform the categorisation and their estimation procedures; (3) the loop status categorisation procedure, and some implementation issues, such as filtered version extensions.

Loop status then forms the basis for the development of a fault localisation technique based on the Overall Loop Performance Index (OLPI). This approach is able to find the source of a plant-wide disturbance, when a number of interacting PI/PID control loops are perturbed by a single dominant disturbance/fault. This work is described in Chapter 4.

The OLPI-based technique is not suitable for localisation/isolation in a tightly coupled plant, which would result in poor performance during the commissioning-stage. For instance, two or more tightly coupled control loops can act as an oscillator or resonator. Unfortunately detailed knowledge of the plant would be needed to explain the physical phenomena that causes the oscillation or resonance.

The analysis that is discussed in Chapter 5 merely associates the problem with poor plant design. A sequence of tests is suggested to accommodate this special case.

A new application of independent component analysis (ICA), called spectral ICA, is then discussed in Chapter 6. It is shown how spectral ICA has the capability to detect and isolate multiple oscillation sources in a chemical process plant. Its key feature is that it extracts dominant spectrum-like independent components from spectra data, which is derived from the time-domain measurements. Each of independent components has a narrow-band peak that captures the behaviour of one of the oscillation sources.

The conclusions of the work are made and some future directions are advised in the final chapter, Chapter 7.

1.5 Originality

The concept of Loop status, OLPI and spectral ICA and their applications to various processes are all original. The work has been presented at various symposiums and conferences (Xia & Howell 2001a; 2001b; 2002a; 2002b; 2003a), published in the Journal of Process Control (Xia & Howell 2003b), and submitted to Automatica (Xia *et al.* 2003).

Chapter 2

Some Background to Process Monitoring and Fault

Detection & Isolation

This chapter first briefly reviews some general methods for process monitoring and fault detection & isolation (FDI), and then focuses on some specific methods that have had direct influences on the development of the approaches proposed.

2.1 Overview

Since the first monograph on FDI was published by Himmelblau (1978), numerous research activities have resulted in many new approaches. See for example, edited books by Patton *et al.* (1989; 2000); monographs by Mangoubi (1998), Gertler (1998), Chen & Patton (1999), Russell *et al.* (2000) and Simani *et al.* (2003) and recent survey papers (Dash & Venkatasubramanian 2000; Frank 1990; Frank & Koppen-Seliger 1997; Garcia & Frank 1997; Isermann 1984; Isermann 1997; Isermann & Balle 1997).

FDI methods can be classified according to their specific approaches.

- I. Model-based approaches, such as observers, filters, estimators, parity equations & balance equations (Basseville 1997; Berton & Hodouin 2003; Bloch *et al.* 1995; Commault 1999; Dingli *et al.* 1996; Edelmayer *et al.* 1997; Edwards *et al.* 2000; Garcia *et al.* 1995; Goh *et al.* 2002; Hamelin & Sauter 2000; Henry *et al.* 2002; Hofling & Isermann 1996; Keller 1999; Kinnaert 1999; McKenzie *et al.* 1998; Medvedev 1995; Patton & Chen 1997; Patton & Hou 1998; Shen & Hsu 1998; Simani *et al.* 2000; Sun *et al.* 2002; Tan & Edwards 2002; Verde 2001; Weyer *et al.* 2000; Yu & Shields 1996; Zolghadri *et al.* 1996).

- II. Artificial intelligence (AI) based methods, such as neural networks (NN) (Brydon *et al.* 1997; Chan *et al.* 2001; Fuente & Vega 1999; Karpenko & Sepehri 2002; McGhee *et al.* 1997; Rengaswamy & Venkatasubramanian 2000; Vora *et al.* 1997; Weerasinghe *et al.* 1998; Yang *et al.* 2000; Yu *et al.* 1999), expert systems (ES) (Kumamoto *et al.* 1984; Rich & Venkatasubramanian 1987; Shum *et al.* 1988), fuzzy logic (Aubrun *et al.* 1995; Balle 1999; Balle & Fuessel 2000; Garcia *et al.* 2000; Leonhardt & Ayoubi 1997), causal digraphs & signed digraphs (SDG) that are based on qualitative modelling & reasoning (Chen & Howell 2001, 2002; Dong *et al.* 1996; Kyung *et al.* 1997; Mohindra & Clark 1993; Vedam & Venkatasubramanian 1997; Wang *et al.* 1995; Wilcox & Himmelblau 1994), miscellaneous intelligent approaches such as pattern recognition, symbolic machine learning, knowledge-based reasoning and so on (Love & Simaan 1988; Ozyurt *et al.* 1998; Rozier 2001; Schiller *et al.* 2001;

Teo & Gooi 1997; Tse *et al.* 1996) and mixed-AI methods that involve two or more AI approaches (Becraft & Lee 1993; Leung & Romagnoli 2000; Montmain & Gentil 2000; Rakar *et al.* 1999; Ruiz *et al.* 2000; Ruiz *et al.* 2001a, 2001b; Tarifa & Scenna 1997; Zhao *et al.* 1997).

III. Methods that are based on statistics and signal processing (data-driven approaches), for example, multivariate statistical techniques like principal components analysis (PCA), partial least squares (PLS) and so on (Amand *et al.* 2001; Dunia & Qin 1998; Huang *et al.* 2000; Jia *et al.* 1998; Kano *et al.* 2000; Kano *et al.* 2001; Kano *et al.* 2002b; Pranatyasto & Qin 2001; Raich & Cinar 1997; Thornhill *et al.* 2002a; Yoon & MacGregor 2001), wavelet techniques (Alexander & Gor 1998; Daiguji *et al.* 1997; Tsuge *et al.* 2000) or wavelet plus PCA techniques (Fourie & de Vaal 2000; Shao *et al.* 1999), data-driven trend/feature analysis, statistical signal processing, frequency domain analysis, independent component analysis (ICA), information theory and so on (Bakshi & Stephanopoulos 1994; Choudhury *et al.* 2002; Dobson & Thornhill 2002; Forsman & Stratin 1999; Hägglund 1995, 1999; Horch 1999; Kano *et al.* 2002a; Miao & Seborg 1999; Rengaswamy *et al.* 2001a; Thornhill *et al.* 2000, 2001, 2002b, 2003; Thornhill & Hägglund 1997).

IV. Synthetic methods that combine some of the above classified methods, such as wavelet plus NN (Chen *et al.* 1998; Chen *et al.* 1999; Zhao *et al.* 1998), model-based plus statistics (Bogaerts *et al.* 1998; Kinnaert *et al.* 2000), NN plus statistics (Bissessur *et al.* 1999; Leger *et al.* 1998) and other various combinations (Akbarian & Bishnoi 2001; Hofling *et al.* 1995; Liu 1999;

Mylaraswamy & Venkatasubramanian 1997; Ouali *et al.* 1999; Rengaswamy & Venkatasubramanian 1995; Vedam & Venkatasubramanian 1999; Zhang *et al.* 1998).

Each of these FDI approaches has clear strengths and weaknesses. For example, the majority of model-based approaches rely on completely known, linear models. Precise models of this type are not available in many cases, particularly in the chemical process industry (Rengaswamy *et al.* 2001b). Data-driven & AI-based methods usually avoid the need for complex quantitative plant modelling, making them more suitable for cases where the plant model is complex and not easy to obtain. Expert system & fuzzy inference are knowledge based heuristic solutions that are simple and understandable. However, the bottle-neck of knowledge acquisition, especially automated in-depth knowledge acquisition in a large-scale plant situation, is not necessarily easy to overcome. Qualitative reasoning, such as SDG or cause-effective digraph based approaches, relies largely on the development of qualitative cause-effective AI models. Such qualitative models are relatively easy to acquire, but the ambiguities associated with these qualitative representations often make the diagnostic performance of these methods quite poor, especially during transients. Thus qualitative reasoning is normally only suitable for monitoring of plants that are in steady state. Neural networks (NN) provide an alternative for where the precise model is not available. The main drawback of NN-based methods is the need for a large amount of typical training data that can reflect different kinds of possible faults. Such training data may not be easy to acquire in chemical process plants, especially historical data pertaining to known fault conditions.

Data-driven based FDI methods have been receiving considerable attention in recent years due to the various drawbacks of model-based & AI-based methods discussed above. It is, however, generally accepted that no single diagnostic method adequately addresses all the challenges of complex, plant-wide diagnostic problems. A hybrid framework combining different properly designed methods for specific plant situations represents a better solution.

The approaches that are proposed in this thesis are mainly data-driven, and thus the emphasis of the rest of this chapter is on the most relevant data-driven methods.

2.2 Control Loop Performance Assessment

Over the past decade or so there has been considerable commercial and academic interest in developing methods for analysing the performance of univariate and multivariate control systems based on the minimum variance control (MVC) benchmark. A comprehensive approach for controller performance monitoring usually includes the following elements:

- (1) the determination of the capability of the control system,
- (2) the development of suitable statistics for monitoring the performance of the existing system,
- (3) the development of methods for diagnosing the underlying causes for changes in the performance of the control system (Harris & Seppala 2001).

Reviews and critical analysis of several approaches for assessing control loop performance can be found in Harris & Seppala (2001), Huang & Shah (1999), Harris *et al.* (1999) and Qin (1998).

2.2.1 MVC Benchmark Based Approaches

The single-input single-output minimum variance control based performance benchmark was the initial underlying concept, and was first proposed by Harris (1989). In his work Harris proposed the use of a minimum variance controller as a lower bound to assess the performance of single loop controllers. The lower bound is estimated from closed-loop operating data and accounts for the process time delay. No other process model information is needed.

Consider a single-input, single-output (SISO) process under regulatory control as shown in Figure 2-1, where Q is the controller, q^{-1} is the back-shift operator, d is the time delay, \tilde{T} is the delay-free plant transfer function, N is the disturbance transfer function and a_t is white noise.

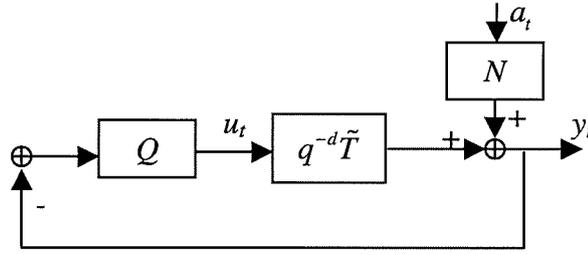


Figure 2-1: Schematic diagram of SISO process under feedback control

The transfer function between the driving force a_t and controlled variable can be given by:

$$y_t = \frac{N}{1 + q^{-d}\tilde{T}Q} a_t \quad (2.1)$$

and N can be decomposed into two parts according to the Diophantine identity:

$$N = \underbrace{f_0 + f_1 q^{-1} + \dots + f_{d-1} q^{-d+1}}_F + R q^{-d} \quad (2.2)$$

where $f_i (i=1 \dots d-1)$ are constant coefficients, and R is the remaining rational proper transfer function. Thus Equation (2.1) can be rewritten as

$$y_t = \frac{F + R q^{-d}}{1 + q^{-d}\tilde{T}Q} a_t = \left[F + \frac{R - F\tilde{T}Q}{1 + q^{-d}\tilde{T}Q} q^{-d} \right] a_t = \underbrace{F a_t}_{e_{mv}} + \underbrace{L a_{t-d}}_{\hat{y}_t} \quad (2.3)$$

i.e.,

$$y_t = e_{mv} + \hat{y}_t \quad (2.4)$$

where $L = \frac{R - F\tilde{T}Q}{1 + q^{-d}\tilde{T}Q}$ is a proper transfer function, e_{mv} is called the minimum variance output independent of feedback control, and \hat{y}_t represents the d -step ahead predictable components. It can be seen that e_{mv} and \hat{y}_t are independent because e_{mv} contains only $a_t \cdots a_{t-d+1}$ terms whilst \hat{y}_t contains terms that are ahead of a_{t-d} , and as a result,

$$\sigma_{y_t}^2 = \sigma_{e_{mv}}^2 + \sigma_{\hat{y}_t}^2$$

Therefore

$$\sigma_{y_t}^2 \geq \sigma_{e_{mv}}^2$$

The equation holds when there are no d -step ahead predictable components, i.e., $\hat{y}_t = 0$, or equivalently $L = 0$, i.e., $R - F\tilde{T}Q = 0$. This yields the minimum variance control law: $Q = \frac{R}{F\tilde{T}}$.

The minimum variance of e_{mv} is a natural lower bound benchmark for the assessment of loop performance, a well performed loop should have little d -step ahead predictable components, and thus $\sigma_{y_t}^2 \cong \sigma_{e_{mv}}^2$, whilst for a poorly turned loop, $\sigma_{y_t}^2 \gg \sigma_{e_{mv}}^2$. The control loop performance index can be defined as

$$\eta_y = \frac{\sigma_{y_t}^2}{\sigma_{e_{mv}}^2} \quad (2.5)$$

or alternatively as

$$\eta_y = \frac{\sigma_{e_{mv}}^2}{\sigma_{y_t}^2} \quad (2.6)$$

or as

$$\eta_y = 1 - \frac{\sigma_{e_{mv}}^2}{\sigma_{y_t}^2} \quad (2.7)$$

Definitions (2.6) and (2.7) are normalised indices, i.e. their values are within [0,1].

A fast, simple, on-line method for estimating index η_y has been proposed by Desborough & Harris (1992), which is based on a d -step ahead prediction method for the estimation of predictable components \hat{y}_t . An m -order linear AR data model can be used to extract \hat{y}_t from the normalised closed-loop sample data $\{y_1, y_2 \dots y_n\}$. The matrix notation of the AR model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{e} \quad (2.8)$$

where

$$\mathbf{y} = \begin{bmatrix} y_n \\ y_{n-1} \\ \vdots \\ y_{m+d} \end{bmatrix}, \mathbf{X} = \begin{bmatrix} y_{n-d} & y_{n-d-1} & \cdots & y_{n-d-m+1} \\ y_{n-d-1} & y_{n-d-2} & \cdots & y_{n-d-m} \\ \vdots & \vdots & \vdots & \vdots \\ y_m & y_{m-1} & \cdots & y_1 \end{bmatrix}, \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$$

$\boldsymbol{\alpha}$ contains the model coefficients that need to be estimated and \mathbf{e} is the residual.

The estimated model coefficients can be obtained by linear regression, i.e.

$$\hat{\boldsymbol{\alpha}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \quad (2.9)$$

the estimated d -step ahead predictable components is then

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\alpha}} \quad (2.10)$$

and the residual, i.e. the estimated minimum variance output, is

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} \quad (2.11)$$

Thus the estimated minimum variance can be obtained from the mean square error of \mathbf{e} , i.e. $\text{mse}(\mathbf{e})$ and then the performance index can be determined. For example, the index defined by equation (2.5) can be obtained by

$$\eta_y = \frac{mse(\mathbf{y})}{mse(\mathbf{e})} = \frac{mse(\mathbf{y})}{mse(\mathbf{y} - \hat{\mathbf{y}})} \quad (2.12)$$

Thornhill *et al.* (1999) have discussed the application of this type of control loop performance assessment (CLPA) to a refinery. An AR data model is utilised in their CLPA algorithm, and recommended CLPA parameters, as shown in Table 2-1, are given for generic refinery loops.

Loop type	Sampling interval (s)	Prediction horizon d (steps)	Data window length n	AR model order
Pressure	20	5	1000-1500	30
Liquid flow	10	3		
Temperature	60-120	5-6		
Steam/gas flow	60	5		
Level	20	5		

Table 2-1: Recommended CLPA parameters

Note that the prediction horizon d is treated as an engineering criterion rather than being derived from the process time delay. The reason for this is as follows:

- (1) it is time consuming and costly to determine the time delay of a control loop in an operating plant because the loop has to be taken off-line for step tests;
- (2) it is found that the precise time delay information is not needed if the choice of prediction horizon is sensible. Thornhill *et al.* (1999) have explored the effects of choosing different prediction horizons d on the assessment of representative loops. For example, the CLPA index vs. d plot in Figure 2-2 can be obtained from normal history operating data for different loops, and a good choice for the prediction horizon is one that falls on the plateau where the CLPA index does not vary rapidly.

Recommended values of d for typical refinery loops are given in Table 2-1.

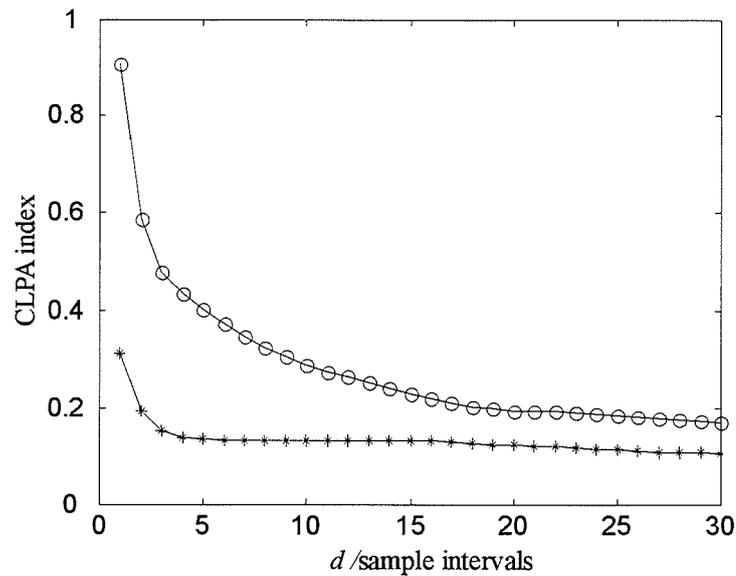


Figure 2-2: CLPA index plot with respect to prediction horizon d

Huang & Shah (1999) have proposed an alternative algorithm, the FCOR algorithm, to estimate the CLPA index. Their method is based on a whitening filter (MA model) to estimate the innovation white noise sequence a_t , and the cross-correlation between y_t and a_t for lags 0 to $d-1$. Lynch and Dumont (1996) have used a Laguerre series model rather than an AR or MA model, and have also proposed a time-delay estimator.

The development of minimum variance benchmark based, off-line closed loop performance assessment (CLPA) techniques are now so well established that various vendors are offering commercial analysis products based on them. For example, there is the *Loop ScoutTM* software package from Honeywell and *ProcessDoctor* from Matrikon. Various industrial experiences with CLPA

applications have been published (Ko & Edgar 2001b; Kozub 1997; Owen *et al.* 1996; Paulonis & Cox 2003; Thornhill *et al.* 1999).

2.2.2 Modifications, Extensions and Other Non-MVC Based Approaches

Many modifications and extensions have been made to the basic MVC-based control loop performance assessment methods, for example, the methods have been extended to feedforward/feedback systems (Desborough & Harris 1993; Stanfelj *et al.* 1993), cascade systems (Ko & Edgar 2000), linear time variant processes (Huang 2002) and been modified for different situations (Bezergianni & Georgakis 2000; Grimble 2002; Horch & Isaksson 1999; Ko & Edgar 2001a). The extension of MVC-based approaches to multivariable systems has been studied by Harris *et al.* (1996), Huang *et al.* (1997), Huang & Shah (1999) and Ko & Edgar (2001b).

Some non-MVC based approaches are worth mentioning briefly here: frequency domain techniques (Kendra & Cinar 1997), likelihood methods (Tyler & Morari 1996) and multivariate statistics based process monitoring techniques, such as PCA and PLS. The latter have been studied by several academic and industrial engineers (See for example, Kano *et al.* 2000; Kano *et al.* 2001; Kano *et al.* 2002b; Lennox *et al.* 2000; MacGregor & Kourti 1995; Martin *et al.* 1996; Russell *et al.* 2000; Simoglou *et al.* 2002). New trends on dimension reduction and process monitoring based on independent component analysis (ICA) were first proposed by Li & Wang (2002) and Kano *et al.* (2002a).

2.2.3 Fault Detection Based on Performance Index

The feasibility of using a performance index as an indicator for fault detection has also been examined. Fasolo and Seborg (1995), who have proposed a fault detection strategy based on the performance index in a simulated HVAC control system, claim that the index can distinguish between faults and the effects of a load disturbance. Thornhill and Hägglund (1997) have proposed an off-line method to detect and diagnose oscillations in control loops. They use several grey indicators such as the performance index (CLPA), the regularity factor, harmonics in the power spectrum, and the sp-pv map for cascade loops. Owen *et al.* (1996) has described how automatic monitoring of control loop performance could be applied to a mill prototype for the detection and location of malfunctioning loops. A modified index was used in his method, along with the Harris index, to indicate the most probable malfunctioning loop within a group of interacting loops.

2.2.4 Criticisms

MVC-based control loop performance assessment techniques in SISO systems are now well established. Commercial analysis products are available, because of the simplicity of the concept and ease of implementation. However, it has been pointed out that the use of the Harris performance index, as the sole criterion for flagging a malfunctioning control loop, is unlikely to locate the malfunction reliably because certain non-linearities or temporary deterministic drifts can either inflate or underestimate the index (Owen *et al.* 1996). Besides the method does not have a fault isolation capability and the Harris index would not even detect a fault/change

in some cases, for example, a well-tuned controller would effectively compensate for a sensor drift or a low frequency disturbance, and thus such fault/changes will not be indicated by changes in the performance index. Further analysis of the root cause of poor performance is critical. Current research efforts in this area have focused on the detection and diagnosis of oscillations in either SISO systems or multiple-loop plant arrangements, and this will be discussed in the next section.

Multivariate performance assessment techniques are yet to be studied fully. Current methods need knowledge of the unitary interactor matrix. Both the concept and the calculation of the interactor matrix have been obstacles for practising control engineers. Unlike the time-delay term in SISO processes, the interactor matrix, in general, can not be constructed from knowledge of time-delays only, and its estimation requires the insertion of a dither signal into the process during the period of data collection. Recently, Ko and Edgar (2001b) proposed a new procedure for the estimation of multivariate minimum variance performance bounds, which does not require knowledge of the interactor matrix, but still needs estimates of the first few Markov parameters (up to the delay order). It does not really reduce the *a priori* knowledge requirement for the calculation of performance index. The diagnosis of the cause of poor performance of multivariable systems has not been investigated much at all.

2.3 Oscillation Detection & Diagnosis

One of the key elements of a comprehensive controller performance monitoring approach must be methods for diagnosing the underlying causes for changes in the performance of the control system. Severe deterioration of performance in process control loops is often indicated by oscillations and progress has been made in the detection and diagnosis of oscillations. It is reported that about 30% of all control loops in a paper mill plant were oscillating because of valve problems (Hägglund 1995). Ender (1993) reported similar results. The oscillations may cause increased energy consumption, waste of raw material and sometimes a less uniform product. Reducing or removing such oscillations yields commercial benefits (Hägglund 1995; Martin *et al.* 1991; Shinskey 1990; Thornhill and Hägglund 1997). Thus simple and efficient methods for oscillation detection and diagnosis is crucial.

2.3.1 Common Causes of Oscillations

There are several possible causes of an oscillation. Control valve non-linearity, such as friction, dead band, backlash or hysteresis, is one of the most possible causes. The fundamental frequency of a limit cycle caused by a valve non-linearity can be estimated by employing a technique based on describing functions. A significant phase shift in the describing function would imply a lower-frequency limit cycle, on the other hand, little phase shift in the describing function would imply a higher-frequency limit cycle, one whose frequency is closer to the process natural frequency. Another possible cause is bad tuning, which can destabilise the

system. An external oscillatory disturbance is the third possible reason. This situation often happens in a plant that contains loop interactions. The source oscillation will propagate to other interacting loops and result in secondary oscillations. Since interaction is inevitable in process plants, the isolation of the oscillation source becomes difficult.

2.3.2 Single-Loop Based Approaches

Oscillation detection based on zero-crossings analysis

One of the popular oscillation detection techniques in SISO systems is based on the analysis of zero-crossings of the loop operating data. Hägglund (1995) presented a real-time oscillation detection method that calculates the integrated absolute deviation (IAE) between successive zero crossings of the controller error signal. The idea is that when the controller error is oscillatory rather than random, such deviations are large and the intervals between them is large, i.e., an oscillation signal has larger IAE values than a random one. The definition of IAE is given by

$$IAE_i = \int_{t_i}^{t_{i+1}} |Y(t)| dt \quad (2.13)$$

where $Y(t)$ is the controller error signal and t_i and t_{i+1} are times of successive zero crossings of $Y(t)$. These IAE deviations are compared to a threshold value which is based on the IAE value of a sinusoidal oscillation of amplitude a and on the highest frequency that might occur in the loop, i.e. the ultimate natural frequency ω_u . The IAE for such a sine wave ($a \sin(\omega_u t)$) is $\frac{2a}{\omega_u}$. When the ultimate frequency is not

known, a good approximation for ω_u could be $\frac{2\pi}{T_i}$, where T_i is the controller integral constant. This is because, in a well-tuned controller, the ultimate natural frequency at which a resonant oscillation occurs is similar to $\frac{2\pi}{T_i}$. Thus the condition for oscillation detection is:

$$IAE_i \geq \frac{2a}{\omega_u} = \frac{aT_i}{\pi} \quad (2.14)$$

In the real-time implementation an oscillation is detected if the IAE deviation exceeds the threshold, with a set to one percent of the controller range over a supervision time of 50 times the presumed oscillation period.

Thornhill (1997) extended Hägglund's zero-crossings idea to an off-line analysis for oscillation detection, and also presented operational signatures that indicated the cause of an oscillation. Forsman and Stratin (1999) improved the zero-crossings techniques to accommodate the detection of asymmetric and irregular oscillations.

Criticism 1

These methods are simple and easy to implement, however excessive noise will degrade their performance. Also the first two techniques are designed under the assumption of symmetric oscillations.

Oscillation detection based on auto-correlation

Miao and Seborg (1999) have proposed a statistical-based approach to the detection of excessively oscillatory feedback control loops. The test is based on a decay ratio, which is obtained from the auto-correlation of the operating data, rather than from the oscillatory operating data itself. The underlying basis is that the auto-correlation

of an oscillatory signal is also oscillatory. Thus if the decay ratio is greater than a specified threshold, it is concluded that the signal is excessively oscillatory. Thornhill *et al.* (2003) proposed a new method that applied the zero crossings analysis to the filtered auto-covariance for the detection of multiple plant-wide oscillations.

Criticism 2

This auto-correlation or auto-covariance based approach is an improved detection method compared to time domain zero-crossing's methods in the sense that the calculation of auto-correlation essentially reduces the noise effect. But the computation burden will be heavier, making it less suitable for online implementation. The application of zero-crossing analysis to auto-covariance data combines the advantages of both and is therefore superior to other off-line detection methods.

Diagnosis of an oscillation

The diagnosis of the root cause of an oscillation in a SISO control loop is largely focused on the analysis and identification of a valve non-linearity. Horch (1999) presented a new method to identify stiction in control valves. The method is based on the cross-correlation between control input and process output. The motivation is as follows: the limit cycle caused by a sticking valve will result in a square wave profile of the controller variable $y(t)$ and a triangular wave profile of the PI controller output $u(t)$ (See Figure 2-3) with a $\frac{\pi}{2}$ phase shift between them, provided that the process has no integral action.

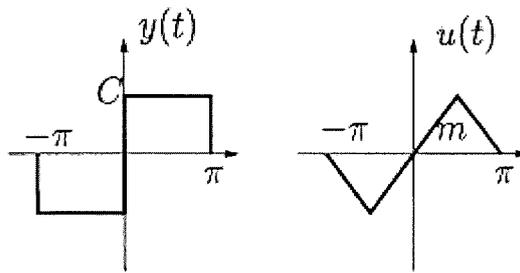


Figure 2-3: Ideal $y(t)$ and $u(t)$ signals in the presence of stiction

The detection condition is then:

- (1) in the case of static friction (stiction) in a control valve, the $y(t)$ and $u(t)$ signals are shifted by approximately $\frac{\pi}{2}$ so that their cross-correlation is an odd function;
- (2) oscillating external disturbances result in a phase shift of approximately π , so that the cross-correlation is an even function. This is because, in general, low frequency disturbances will be eliminated efficiently by the PI-controller since a controller with integral action has a high loop gain at low frequencies;
- (3) an unstable loop which oscillates with constant amplitude (due to physical saturation) also results in an even cross-correlation function.

Both conditions (2) and (3) imply a high frequency oscillation such that the phase shift between $y(t)$ and $u(t)$, which is caused by the PI controller, is small and the only phase shift, i.e. π , is caused by the negative feedback sign.

A similar shape-based idea was used by Rengaswamy *et al.* (2001a) to detect and diagnose different kinds of oscillations in control loops. The method is based on a semi-qualitative approach for the identification of different shapes of oscillation:

square, triangular and non-triangular. The diagnosis of valve friction or hysteresis can then be achieved by examining the shape of the oscillation of the controlled variable and the controller output. For example: (1) a triangular controller output plus a non-triangular controlled variable corresponds to a situation of asymmetric friction; (2) a triangular controller output plus a square controlled variable corresponds to the case of hysteresis.

Taha (1996) has presented an automatic procedure for the diagnosis of oscillations. The focus is on the detection of valve friction. By analysing the friction degree, a sticking valve can be identified. A process-model based approach is also used to distinguish between a poorly tuned controller or the presence of an external oscillatory disturbance, when it is not caused by valve friction.

Thornhill *et al.*(2001) have introduced a non-linear time series analysis approach to detect the existence of non-linear components in control loops. The method compares the prediction error of the test time series (i.e. the process variable) with a statistical prediction error value calculated from a set of surrogate data, which has the same statistical characteristics up to second order (e.g. its power spectrum or auto-correlation function). The prediction error will be similar to that of surrogate data if it is linear, otherwise it is non-linear.

Choudhury *et al.* (2002) have proposed non-linearity test methods that are based on the fact that, in higher order statistics, the squared bicoherence is constant for a linear process. If the bicoherence is not a non-zero constant, it is concluded that the

process is non-linear. After a non-linear system is detected, a *pv-op* map is used to associate the non-linearity with one of a set of known non-linear characteristics.

Criticism 3

Horch's sticking valve detection method is simple and straightforward, and process model information is not required. However it has some drawbacks: (1) the distortion of the ideal $y(t)$ and $u(t)$ signal shapes in the presence of stiction could degrade the detection performance; (2) a low frequency external oscillatory disturbance could still be dominant in the perturbed loop, even if it is largely attenuated, and such a low frequency oscillatory disturbance will result in an odd cross-correlation function, which will contradict the second detection condition.

Rengaswamy *et al.*'s method largely relies on the oscillation shape. A low degree of valve non-linearity may not cause the typical pattern, thus the diagnosis of the oscillation might not be effective in such cases. However, the oscillation detection algorithm has quite good performance, because it can recover shapes from data with up to 20% noise.

Taha's method is a model based approach because it needs valve input-output data and the valve characteristic equation. Such model valve input-output data information, in general, is not easy to obtain, and this has prevented such methods from being widely accepted by plant practitioners. Data-driven approaches would be more attractive to process engineers.

Data-driven approaches based on non-linearity testing (either non-linear time series analysis or higher order statistics) are relatively new and should attract more research activity. The described two approaches, however, have a common drawback in that the control system is assumed to have only white noise or independent identically distributed (i.i.d.) noise: this assumption does not hold in the situation of plant-wide oscillations, where most loops contain external oscillatory disturbances, which might exhibit non-linear behaviours. It is anticipated that false alarms might be generated. It is therefore better to isolate the oscillation source first, and then do the non-linear test on that source loop to reduce the risk of false detection and to minimise the computational effort. Note that the localisation & isolation of plant-wide oscillation sources is one of the main contributions of this thesis.

2.3.3 Analysis of Plant-Wide Oscillations

All the oscillation detection and diagnosis methods discussed above have focused on a single SISO control loop. This section will discuss methods available for the detection and clustering of multiple plant-wide oscillations.

Thornhill *et al.* (2002a) have described the principal component analysis (PCA) of the power spectra of data from chemical processes. The novel feature of spectral PCA is that the rows of the data matrix, X , are the single-sided power spectra $P(f)$ of the signals over a range of frequencies up to the Nyquist frequency:

N frequency channels \rightarrow

$$\mathbf{X} = \begin{pmatrix} P_1(f_1) & \cdots & P_1(f_N) \\ \cdots & \cdots & \cdots \\ P_m(f_1) & \cdots & P_m(f_N) \end{pmatrix} \quad \begin{array}{l} m \text{ process variables} \\ \downarrow \end{array}$$

By performing PCA of this data, the power spectra in \mathbf{X} can be decomposed into combinations of several dominant spectrum-like principal components (PCs):

$$\mathbf{X} = \begin{bmatrix} t_{1,1} \\ t_{2,1} \\ \vdots \\ t_{m,1} \end{bmatrix} \mathbf{v}'_1 + \begin{bmatrix} t_{1,2} \\ t_{2,2} \\ \vdots \\ t_{m,2} \end{bmatrix} \mathbf{v}'_2 + \begin{bmatrix} t_{1,3} \\ t_{2,3} \\ \vdots \\ t_{m,3} \end{bmatrix} \mathbf{v}'_3 + \mathbf{E}$$

where \mathbf{v}'_1 to \mathbf{v}'_3 represent the dominant row-major spectrum-like PCs, \mathbf{E} is an error matrix, which includes truncated principal components and $t_{i,j}$ (for $i=1\dots m$, $j=1\dots 3$) are mixing factors. For example the i^{th} row power spectrum in \mathbf{X} are approximately constructed by \mathbf{v}'_1 to \mathbf{v}'_3 with the relevant mixing factors $(t_{i,1}, t_{i,2}, t_{i,3})$.

A single dominant peak in a power spectrum, indicating an oscillation in the time domain, can be created by proper linear combinations of the PCs. The clustering of similar oscillatory trends is achieved by a scores plot, in which the point $(t_{i,1}, t_{i,2}, t_{i,3})$ maps to the i^{th} power spectrum. Similar spectra have similar t -coordinates. Therefore such groups form clusters. PCA of the auto-covariance functions of process variables gives similar or sometimes even superior performance compared to spectral PCA.

Thornhill *et al.* (2003) have presented an automated tool for the detection of oscillations in measurements from chemical processes including the case when two

or more oscillations of different frequency are present simultaneously. The presence of oscillations in selected frequency ranges is determined using a new method based on the regularity of the zero crossings of the filtered auto-covariance data. An automated filter selection algorithm for the selection of frequency range is proposed.

Criticism 4

Spectral PCA has superior performance to time domain PCA for the detection of clusters of data trends having similar features, even when time shifting is used to align the data trends. This is because spectra are invariant to the phase lags caused by time delays and process dynamics. Although the interpretation of what is detected, and the grouping of oscillation disturbances based on spectral PCs, is not straightforward because the same spectral peaks could be present in more than one PC and becomes difficult when the number of dominant PCs increases, it is still a good approach to the identification of the dominant spectral peaks. The multiple-oscillation detection algorithm is a good automated tool, which is totally data-driven. However neither of the two methods address the isolation of these oscillations, which is of greater interest to plant operating engineers. The isolation of multiple plant-wide oscillations in measurements from chemical processes has not yet been thoroughly studied.

Chapter 3

Loop Status Monitoring

This chapter describes a new concept, that of *Loop Status* monitoring, which indicates the current deterministic trend of a loop. The chapter emphasises:

- the categorisation of a deterministic trend as one of seven statuses;
- the scientific basis of the statistics needed to perform the categorisation and their estimation procedures;
- the loop status categorisation procedure, and some implementation issues, such as filtered version extensions.

3.1 Overview

Although loop performance monitoring in SISO systems has been well-researched, especially minimum variance benchmark based techniques (Harris *et al.* 1999; Harris & Seppala 2001; Huang & Shah 1999), the emphasis has been on audit, in a sense that the underlying performance is of issue, rather than on what the loop is actually doing at any particular time. When the focus has been on the actual trend itself, workers have looked at the shape of the trend (Bakshi & Stephanopoulos 1994; Janusz & Venkatasubramanian 1991; Rengaswamy *et al.* 2001a; Wong *et al.* 2001) with the aim of diagnosing faults or disturbances rather than of determining loop performance. Control loop status monitoring can be defined as the near-real-time declaration of what kind of deterministic trend a loop is actually exhibiting.

The status might indicate that the loop is (1) well-behaved & in steady state, or (2) well-behaved but with controller compensation, or (3) undergoing a short-term transient, or (4) undergoing a trend that is disturbed in some non-stationary manner, or (5) cycling at a relatively low fundamental frequency, or (6) cycling at a fundamental frequency similar to the natural frequency of the loop, or (7) out of control (Critical). For instance category (4) might arise because of feedstock variability whilst category (5) might arise because of a cyclic disturbance, or of a valve problem involving a phase lag. Transients can be classed as either short-term or long-term depending on whether or not the transient lasts longer than the loop's settling time. This is an indicator of whether the response is caused solely by a short-term event, like a change in operating point or by something more sustained. From a performance point of view the long-term is more important. Some of the proposed loop statuses might be of interest to the operator, whilst others might be of more interest to the control/maintenance engineer. Some might be of use when analysing data off-line. For instance the maintenance engineer would like to know when a loop is limit cycling, whereas someone analysing off-line data would like to know if the plant was operating in an unsteady manner. The benefit might be as much to do with providing information about plant operation, as with understanding the performance of an individual control loop. This would particularly be so, for instance, if the controllers were compensating for a leak.

This chapter proposes various statistics to facilitate the status monitoring of PI/PID loops and to isolate a problem loop. These statistics can be complemented by the control loop performance monitor of Hägglund (1995), which focuses on Statuses (5) & (6), and by the automatic detection of sluggish control loops (Hägglund

1999), which can be performed during Status (3). Although loop status is a qualitative description, a quantitative version can be formulated by assigning different real number values to the various categories.

Some possible causes for these loop statuses are given in Table 3-1. Note that although a loop status could indicate some possible faults or changes in the operation of loop, it might not necessarily contain any fault, because the fault or disturbance might propagate from another interacting loop. The localisation of a single root-cause loop in a number of interacting loops will be discussed in Chapter 4. Isolation of multiple faults/disturbances, especially in a plant-wide oscillation context, will be discussed in Chapter 6. The categorisation of loop status for a control loop is based on several time-domain statistics that are derived from measurements of the controlled variable & the controller output variable. No process model is needed.

Loop Status	Cause
Compensated	a small disturbance, small fault such as sensor null shift or slow drift
Short-term transient	a step load disturbance (loop dynamic change), sensor bias, sensor large drift or short-term large disturbance
Long-term trend	(A) a non-stationary disturbance, feedstock variability, raw material variability (B) cyclic disturbance or valve problem involving a phase lag
Ultimate cyclic (PI)	ultimate-cyclic disturbance or bad tuning, a valve problem without a phase lag
Critical	a critical change in process

Table 3-1: Some possible causes for loop statuses

3.2 Fundamentals

This section describes a set of statistics, which, when examined together, provide a means of deciding the Loop Status of either a single PI/PID loop, or the master loop of a cascade PI/PID-based control system. An explanation is then given of how these statistics might be estimated from plant data.

3.2.1 The Model

Figure 3-1 shows linear block diagram representation of a control loop. Blocks P & H represent the process and controller transfer functions, u_t is the controller output, y_t is the controlled variable and d_t is a deterministic disturbance, which is modulated with additive Gaussian white noise, e_t , that is independent of d_t . Non-linear effects caused by, for instance, hysteresis, are accommodated by representing them as '*pseudo*' external deterministic disturbances. Note that the set-point is omitted i.e. it is assumed that the loop is a regulator. The approach does accommodate infrequent changes to the set-point, and such changes are seen as additional disturbances to d_t . The reason for choosing this model framework, which includes a deterministic disturbance, is that most perturbed loops contain external deterministic disturbances, rather than white noise disturbances, due to loop interaction and fault/disturbance propagation. For example, a whole plant oscillation could originate from one root-cause loop and this source loop will then spread out the oscillatory disturbance to other interacting loops, and all these perturbed loops will have deterministic oscillatory disturbances. A loop should

perform well when it is well tuned and deterministic disturbance d_t is zero or insignificant.

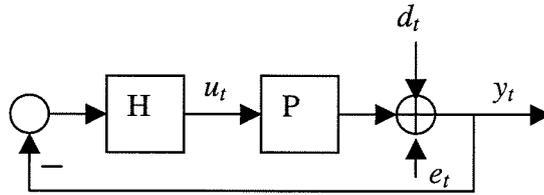


Figure 3-1: Loop with deterministic disturbance & white noise excitation

The status of a loop can be described by the deterministic, or underlying, trend that is observed. The trend of interest is therefore the deterministic component of time series y_t . If \hat{y} is defined as the deterministic component of y_t , and e_y is defined as the direct effect of the noise on y_t , then by definition

$$y_t = \hat{y} + e_y = \frac{1}{1+HP} [d_t + e_t] \quad (3.1)$$

which can be separated into deterministic and stochastic components:

$$\hat{y} = \frac{1}{1+HP} d_t, e_y = \frac{1}{1+HP} e_t \quad (3.2)$$

Similarly the controller output time series can be represented by

$$u_t = \underbrace{(-H\hat{y})}_{\hat{u}} + \underbrace{(-He_y)}_{e_u} \equiv \hat{u} + e_u \quad (3.3)$$

where \hat{u} is the deterministic time series observed in the controller output and e_u is the noise.

Note that e_y is filtered white noise. Figure 3-2 shows the frequency response of this filter, i.e., the sensitivity function, for a typical PI control loop with first order process plant with time delay. Note that it has a high-pass filtering feature, with a

cut-off frequency B_0 , so that e_y will nearly behave as banded white noise with lower cut-off B_0 and upper cut-off at the *Nyquist* frequency $B : B = \left(\frac{\pi}{T_s} \right) \text{rads}^{-1}$, where T_s is the sample interval. The precise location of B_0 is not important, as will be discussed later.

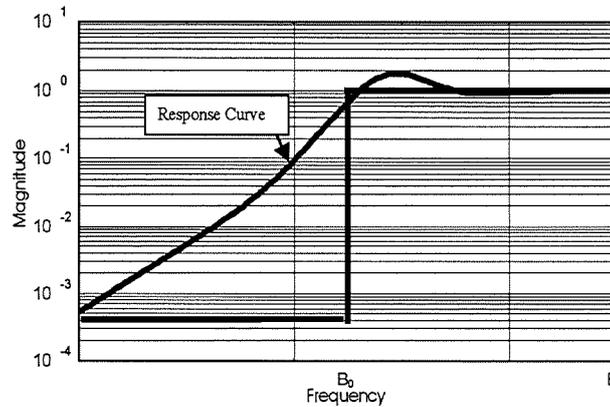


Figure 3-2: Filtered white noise frequency response for a typical PI control loop

3.2.2 The Basis

Sinusoidal trends are common in control loops and can be categorised as either Loop Status (2) compensating (if the frequency is very low) or (5) - (6) cycling (if the frequency is either relatively low or near the natural frequency of the loop). An examination of the behaviour of a PI/PID control loop, when subjected to a pure sinusoidal disturbance of frequency ω , therefore provides a basis for understanding the approach. This behaviour is assessed by analysing three loop statistics, η_y , η_u and their ratio R . The behaviour of a normalised version of R , called R_n , is also examined, particularly in relation to its ability to contribute to the categorisation of trends with other Loop Statuses. In particular, its standard deviation, σ_{R_n} , is found to provide discrimination between Statuses (4) and (5). It will be shown that the

analysis of statistics η_y , η_u , R_n and σ_{R_n} combined, provide the necessary discrimination between the various statuses of a PI/PID loop. Differences in their behaviour between PI & PID control are also examined.

Definition 3-1: Signal-to-noise ratio indices η_y and η_u are defined as:

$$\eta_y = \frac{\sigma_{\hat{y}}^2}{\sigma_{e_y}^2}, \quad (3.4)$$

$$\eta_u = \frac{\sigma_{\hat{u}}^2}{\sigma_{e_u}^2}, \quad (3.5)$$

where $\sigma_{\hat{y}}^2$, $\sigma_{e_y}^2$, $\sigma_{\hat{u}}^2$ and $\sigma_{e_u}^2$ are the variances of \hat{y} , e_y , \hat{u} and e_u respectively.

Statistic η_y will be the same as the minimum variance performance index (Harris *et al.* 1999; Huang and Shah 1999) when the predicted, deterministic component (called the *predictable* component) is estimated by a d -step ahead AR model, where d is the time delay. Although in theory a perfect control loop operating in steady state will have an η_y of zero, industry practice is to recognise that small disturbances do arise and to deem a loop to be performing well if $\eta_y < \gamma$, where γ is typically 0.5 for a temperature loop, but might be smaller for other loops, such as flow loops. A customised value could be obtained by analysing the normal steady state operating data.

Definition 3-2: R is defined as the ratio of the indices:

$$R = \frac{\eta_y}{\eta_u} \quad (3.6)$$

The approach here is first to consider the assessment of a loop, which is performing cyclically, and then to look at the other statuses. With reference to Section 3.2.1, if d_t is a sustained cyclic disturbance with a specific frequency ω and constant magnitude, then after any transient response decays, the long-term response of \hat{y} will also be cyclic, of the same frequency ω , and with a constant magnitude

$$A = \left| \frac{1}{1+HP} d_t \right|. \text{ The next theorem assesses such long-term sustained trends.}$$

Theorem 3-1: If a control loop is under PI control of form $K_p \left[1 + \frac{1}{T_i s} \right]$, where K_p

is the gain and T_i is the integral time constant, and the process has a sine wave oscillatory trend of frequency ω and magnitude A , which is modulated by banded noise e_y with uniform power spectrum of amplitude G over the bandwidth $[B_0, B]$, then

$$\eta_y = \frac{\frac{1}{2} A^2}{G(B - B_0)} \quad (3.7)$$

$$\eta_u = \underbrace{\left(\frac{1 + \frac{1}{(T_i \omega)^2}}{Q} \right)}_{J(\omega)} \eta_y = J(\omega) \eta_y, \quad (3.8)$$

$$R(\omega) = \frac{1}{J(\omega)} = \frac{Q}{1 + \frac{1}{(T_i \omega)^2}}, \quad (3.9)$$

where $Q = 1 + \frac{K_B}{\pi^2} \left[\frac{T_s}{T_i} \right]^2$, T_s is the sampling time and $K_B = \frac{B}{B_0}$.

Proof: See the A.1 Appendix.

Thus there is a simple relationship between η_y and η_u , which can be represented by a meaningful frequency diagram. Also η_y is independent of frequency and constant for a sinusoid of fixed magnitude and stationary noise. The assumption of a uniform spectrum is justified, because the spectrum magnitude G has the same affect on both η_y and η_u and hence doesn't feature in the final result of R in Equation (3.9). This implies that the magnitude of G , and the variance of e_y , i.e. $G(B-B_0)$, will have no influence on R . Note that for most process control loops, $Q \approx 1$, because $\frac{T_s}{T_i} \ll 1$. In fact the value of Q is mainly determined by $\frac{T_s}{T_i}$ rather than K_B (i.e., $\frac{B}{B_0}$).

Table 3-2 shows the sensitivity of changing Q with respect to K_B and $\frac{T_s}{T_i}$. The relative magnitudes of $\frac{\partial Q}{\partial K_B}$ and $\partial Q / \partial [\frac{T_s}{T_i}]$ indicates that a value for K_B , or equivalently a value for B_0 , does not need to be chosen accurately, and the approximated value of Q can be given by: $Q \cong 1 + K \left[\frac{T_s}{T_i} \right]^2$, $K \in [2, 5]$. It implies that even if K_B varies from about 20 to 50 due to the equivalent variation of B_0 , the value of Q only varies from 1.02 to 1.05. An alternative approximate value of Q can be obtained by assigning B_0 the value of $\frac{1}{T_i}$, i.e. the joint point of the low-

frequency asymptote & the high-frequency asymptote of the PI controller, and this

results in $Q \cong 1 + \left(\frac{1}{\pi}\right)\left(\frac{T_s}{T_i}\right)$. For the value of $\frac{T_s}{T_i} = \frac{1}{10}$, $Q = 1.03$.

$\frac{\partial Q}{\partial K_B}$	$\partial Q / \partial \left[\frac{T_s}{T_i}\right]$	When $\frac{T_s}{T_i} = \frac{1}{10}$, $K_B = 20$
0.001	0.4	

Table 3-2: Sensitivity analysis of Q

Note from Equation (3.7) that η_y is frequency independent. Thus η_y will remain constant as long as the magnitude A and the noise properties remain unchanged. Equation (3.8) then shows that η_u can be viewed as η_y 'amplified' by the frequency-related factor $J(\omega)$. At low frequencies $\eta_u \gg \eta_y$, but η_u converges on η_y as the frequency is increased. An example of this is given in Figure 3-3, which shows a typical frequency plot of the three key statistics η_y , η_u & R for a well-performing control loop with $Q=1$, $T_i=20$, and $\eta_y=0.5$. For a well performing loop with a constant η_y , ratio η_u behaves quite differently at low frequencies in Figure 3-3. Consider the case when η_u exceeds γ , whilst at the same time η_y does not (i.e. $\eta_u > \gamma$ & $\eta_y < \gamma$). Here there is a significant predictable signal in the controller output and very little in the controlled variable. This implies that the controller is having to compensate to regulate the controlled variable to set-point. Loop Status (2) denotes this situation. Incidentally this result confirms ones intuitive belief that a low

frequency trend will be more visible in the controller output than in the controlled variable because the controller output will have a higher signal to noise ratio.

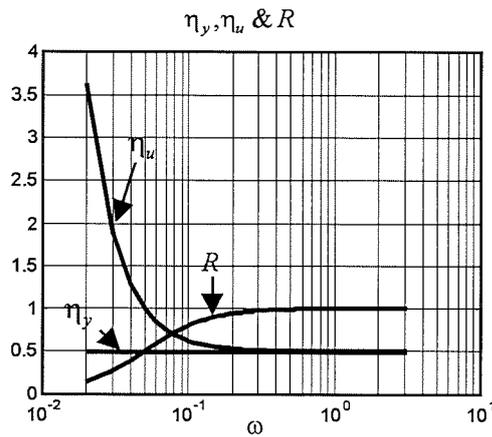


Figure 3-3: η_u , R relationships with ω when $\eta_y=0.5$

In theory, there should be no underlying trends in a well-performing loop operating in steady state, so the well-established performance indicator η_y should be smaller than 0.5. Its detection threshold γ is set relatively large because in practice there might be underlying low frequency components present, which are of no interest to a loop audit. However such trends might be of interest to the operator. If η_u is now compared to the same threshold as η_y , then the same trends will also be 'invisible' to the η_u test unless they are of sufficiently low frequency. If the loop, itself, is performing badly, then both η_y & η_u will exceed their thresholds. In this case another statistic, ratio R , is needed to detect a predominately low frequency trend because R obtains a value $\ll 1$ when there are low frequency trends present. A normalised version R_n is preferred for discriminating between types of oscillation, because this will make the formulation of a test more straightforward.

Remark 3-1: A normalised value of R is defined as:

$$R_n = \frac{R}{Q} = \frac{1}{1 + \frac{1}{(T_i \omega)^2}} \quad (3.10)$$

Figure 3-4 shows a plot of R_n versus $T_i \omega$ for both theoretical and simulation results.

It is significant that $R_n \geq 0.8$ when $T_i \omega \geq 2$ (or $\omega \geq \frac{2}{T_i}$) because this can be used to

discriminate between low and ultimate frequency trends.

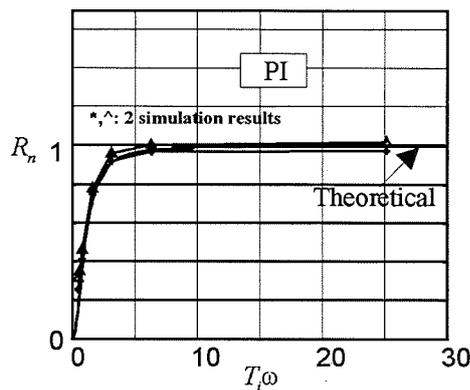


Figure 3-4: Plots of R_n versus $T_i \omega$ for PI controllers

Remark 3-2: The $R_n - \omega$ relationship of Equation (3.10) can be extended to more general deterministic oscillatory trends such as the non-sinusoidal periodic oscillations typical of limit cycles, where ω is now the dominant, fundamental angular frequency of the trend, and the result of $R_n \geq 0.8$, when $\omega \geq \frac{2}{T_i}$, still holds

(See the A.2 Appendix).

Equation (3.10) shows that, for any specific controller, the ratio, R_n , will be solely frequency dependent and is independent of the value of η_y , even if controller performance is poor. This means that R_n can be used to distinguish between certain

oscillatory trends because it obtains a value $\ll 1$ when the trend is of low frequency. Hägglund (1995) has pointed out that, if a PI controller is properly tuned, then usually the integral time constant approximates to the ultimate oscillation period of the closed loop system i.e. the period of oscillation that occurs when its gain is increased. Substituting $\omega = \frac{2\pi}{T_i}$ into Equation (3.10) gives $R_n \approx 1$ i.e.

R_n will be on the plateau of the R_n versus ω plot if a loop is oscillating at its ultimate frequency. This property can be used to discriminate between an ultimate cyclic oscillation and an oscillatory disturbance of a lower frequency. A conservative test can be formulated based on $R_n > R_n \left(\omega = \frac{2}{T_i} \right) = 0.8$, because an ultimate oscillation will have a frequency of about $\frac{2\pi}{T_i} > \frac{2}{T_i}$.

Any variation of R_n with time will suggest that there might be some form of low frequency, non-stationary behaviour, which has a continually shifting dominant frequency. Any non-stationary behaviour at higher frequencies would not have any affect on R_n whereas, as can be seen from Fig. 3, this behaviour at lower frequencies will result in a significant variation in R_n . Such variation can be parameterised by estimating the moving standard deviation, σ_{R_n} , with time: a large σ_{R_n} will indicate low frequency, non-stationary behaviour.

Most of the above also holds for PID control, but with the exception of R , which will be discussed below.

Extension 3-1: If a control loop is under PID control of form $K_P \left[1 + \frac{1}{T_i s} + T_d s \right]$

and the process has an oscillatory trend of frequency ω , then

$$R = \frac{Q_D}{\left(1 - \frac{2}{K_d}\right) + \frac{(T_i \omega)^2}{K_d^2} + \frac{1}{(T_i \omega)^2}} \quad (3.11)$$

where K_d is the ratio of the controller integral time constant T_i to the derivative time constant T_d , and Q_D is a constant as defined in the Appendix A.3.

Proof: See A.3 Appendix.

Remark 3-3: A normalised value of R can be then be defined as:

$$R_n = \frac{R}{Q_D} = \frac{1}{\left(1 - \frac{2}{K_d}\right) + \frac{(T_i \omega)^2}{K_d^2} + \frac{1}{(T_i \omega)^2}} \quad (3.12)$$

The frequency response of R_n for PID action is shown in Figure 3-5, which differs from Figure 3-4 in that there is no high frequency plateau. This derivation is for a particular form of PID controller. The frequency responses of R_n for other forms of PID controller should have a similar shape (some differences will appear in high frequency range), but will have different equations for Q_D . A generic way to calculate the value of Q_D is given in the Appendix A.3.

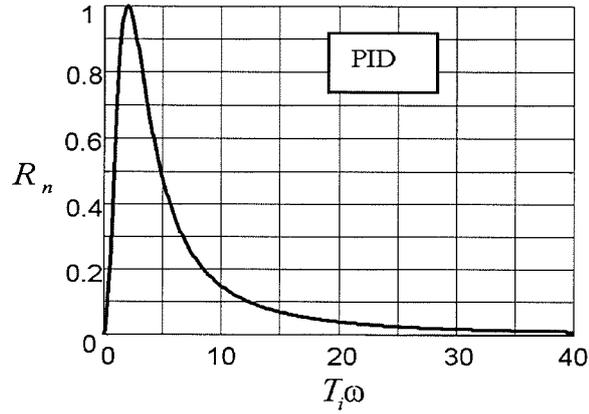


Figure 3-5: Plots of R_n versus $T_i\omega$ for PID controllers

3.2.3 Estimation of η_y , η_u , R_n and Standard Deviation of R_n

A moving data window with N latest sampled data, $\{Y(i), i=1 \cdots N\}$ and $\{U(i), i=1 \cdots N\}$, is used to estimate both η_y and η_u after every sample interval.

Let the estimation results be the time series $\hat{\eta}_y(k)$ and $\hat{\eta}_u(k)$, where k stands for the discrete time sequence.

Step 1: Normalisation.

The data is normalised as follows:

$$y(i) = \frac{Y(i) - \mu_Y}{\sigma_Y}, \quad (3.13)$$

$$u(i) = \frac{U(i) - \mu_U}{\sigma_U}, \quad i=1 \cdots N \quad (3.14)$$

where μ_Y (or μ_U) and σ_Y (or σ_U) are the sample mean and sample standard deviation of Y (or U). For computational simplicity, the controller set-point is used instead of

μ_y when the set-point does not change, because there is little difference between them when a loop is well tuned.

Step 2: Extraction of deterministic trends $\hat{y}(i)$ and $\hat{u}(i)$ from $y(i)$ and $u(i)$.

Many de-noising techniques are available to estimate predictable components from a time series. A 30th order d -step ahead AR model with least squares estimation was adopted in the applications described here, with a data window of length $N=1000$. The choice of prediction horizon d , model order and data window length, was based on the recommendations of Thornhill *et. al.* (1999), who also discuss how these parameters affect the models that are obtained.

Step 3: Estimation of $\hat{\eta}_y(k)$ and $\hat{\eta}_u(k)$

Estimates of the signal-to-noise ratio indices η_y and η_u at time k can then be derived as:

$$\hat{\eta}_y(k) = \frac{\sigma_{\hat{y}}^2}{\sigma_{e_y}^2} \quad (3.15)$$

$$\hat{\eta}_u(k) = \frac{\sigma_{\hat{u}}^2}{\sigma_{e_u}^2} \quad (3.16)$$

where $\sigma_{\hat{y}}^2$ and $\sigma_{e_y}^2$ are the sample variances of \hat{y} and the residual noise,

e_y : $e_y = y - \hat{y}$; $\sigma_{\hat{u}}^2$ and $\sigma_{e_u}^2$ are sample variance of \hat{u} and the residual noise,

e_u : $e_u = u - \hat{u}$.

An estimate of normalised R_n can be obtained at time k from $\hat{\eta}_y(k)$, $\hat{\eta}_u(k)$ and either constant Q or constant Q_D , depending on the type of controller.

For PI controller:

$$\hat{R}_n(k) = \frac{\hat{\eta}_y(k)}{Q\hat{\eta}_u(k)} \quad (3.17)$$

For PID controller:

$$\hat{R}_n(k) = \frac{\hat{\eta}_y(k)}{Q_D\hat{\eta}_u(k)} \quad (3.18)$$

The exponential weighted moving average (EWMA) approach provides an estimate of the sample standard deviation of R_n at time k , i.e., $\sigma_{\hat{R}_n}(k)$. The EWMA of R_n , i.e. $\mu_{\hat{R}_n}(k)$, was first calculated:

$$\mu_{\hat{R}_n}(k) = (1 - \lambda_1)\mu_{\hat{R}_n}(k-1) + \lambda_1\hat{R}_n(k) \quad (3.19)$$

and then

$$\sigma_{\hat{R}_n}^2(k) = (1 - \lambda_2)\sigma_{\hat{R}_n}^2(k-1) + \lambda_2(\hat{R}_n(k) - \mu_{\hat{R}_n})^2 \quad (3.20)$$

where $\lambda_1 \in [0,1]$ and $\lambda_2 \in [0,1]$ are constant factors ($\lambda_1 = \lambda_2 = 0.1$ were used as defaults).

3.3 Loop Status Monitoring

This section first describes how the categories of loop status are determined. Then an extension to quantitative & filtered version of loop status statistic is introduced so as to get a more robust and smooth statistic.

3.3.1 The Qualitative Loop Status Statistic

Tests for PI/PID control, for the various statuses, are given in Table 3-3. Note that categories (4) & (5) are both viewed as 'long-term'. Brief explanations are given below.

Loop Status		Criteria		Default value
Steady (1)		$\hat{\eta}_y < \gamma, \hat{\eta}_u < \gamma$		$\gamma = 0.5$ $\xi = 0.05$ $T_{st} = 24d$ <i>d: time delay</i>
Compensated (2)		$\hat{\eta}_y < \gamma, \hat{\eta}_u \geq \gamma$		
Short-term	Decaying Transient (3)	$\hat{\eta}_y \geq \gamma, \hat{\eta}_u > \gamma, \hat{R}_n \leq 1.2$ $T < T_{st}$		
Long-term	Non-stationary (4)	$\hat{\eta}_y \geq \gamma, \hat{\eta}_u > \gamma$	$\sigma_{\hat{R}_n} > \xi$	
	(5) low-f cyclic (PI) cyclic (PID)	$T > T_{st}$ $\hat{R}_n \leq 0.8(PI)$ $\hat{R}_n \leq 1.2(PID)$	$\sigma_{\hat{R}_n} \leq \xi$	
(Special case for PI) Ultimate cyclic (6)		$\hat{\eta}_y \geq \gamma, \hat{\eta}_u > \gamma$ $0.8 \leq \hat{R}_n \leq 1.2$		
Critical (7)		$\hat{\eta}_y > \gamma, \hat{\eta}_u \geq \gamma, \hat{R}_n \geq 1.2$ <i>or: $\hat{\eta}_y \geq \gamma, \hat{\eta}_u < \gamma$</i>		

Table 3-3: Loop Status criteria

(1) Normal steady operation. Both y_t and u_t should have little predictable component, so the signal-to-noise ratio should be small, i.e., $\eta_y < \gamma$, $\eta_u < \gamma$, where the threshold γ is chosen to accommodate normal acceptable variations whilst operating steadily.

(2) Compensated. A low-frequency disturbance is rejected, effectively, with the net effect that the deterministic trend in y_t is small, whilst the trend in u_t is much larger so that $\eta_y < \gamma$, $\eta_u \geq \gamma$.

All other statuses arise when both y_t and u_t have dominant deterministic predictable components i.e. $\eta_y > \gamma$ and $\eta_u > \gamma$.

(3) Short-term transients. The time interval of interest relates to the loop settling time, T_{st} . If this is known, then a short-term transient is deemed to exist if $\eta_y > \gamma$ & $\eta_u > \gamma$, but only over a period of time $\Delta T : T < T_{st}$. The settling time limit, T_{st} , can be loosely determined by $24d$, where d is the prior knowledge of time delay of the process, if available. Otherwise a customised limit need to be determined.

(4) & (5) Long-term trends. Long-term trends are those trends that last longer than the settling time limit T_{st} . A deterministic, cyclic trend exists if an appropriate threshold ξ can be specified: $\sigma_{R_n} < \xi$. Fig. 5 shows probability density functions (pdfs) of various typical disturbances applied to a simulation of a closed loop system. For a periodical type disturbance, the shape of pdf($\hat{R}_n(k)$) is very narrow, whilst for a long term non-stationary disturbance, the distribution of \hat{R}_n is much

wider and hence the value $\sigma_{\hat{R}_n}(k)$ is much larger. A default value of 0.05 is found to be appropriate for ξ .

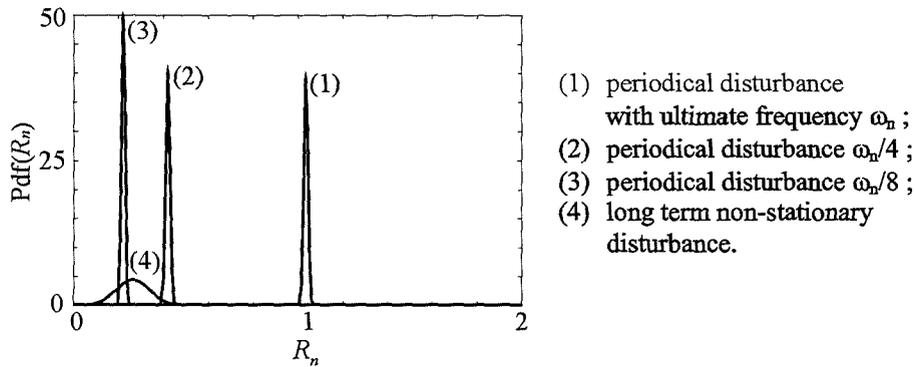


Figure 3-6: Results of $\text{pdf}(\hat{R}_n(k))$

(6) Ultimate cyclic. Occurs when $R_n \approx 1$. A special status only applicable to PI control loops. It implies that the dominant fundamental frequency of a trend is near to nature frequency, i.e., the loop is deemed to be dominated by an ultimate frequency component. This status indicates the possibility of a resonant cyclic that might lead to an unsteady system.

(7) Critical. This category is created to accommodate those test outcomes, which cannot be attributed to the other statuses. These test outcomes can arise when the controller output proves to be ineffective, for instance because the loop is diverging and/or the controller output saturates. In all these cases $\eta_y \geq \gamma$. The value of ratio η_u will depend on the situation at that time. For instance if the controlled variable is diverging whilst the controller attempts to compensate, $\eta_u > \gamma$ but $\eta_y \geq 1.2\eta_u$, i.e., $R_n \geq 1.2$ (this tolerance is sufficiently large to accommodate

uncertainty in the value of Q). Alternatively, the control output might be saturated, and hence constant, in which case η_u is contrived to be zero.

The loop status criteria for PID control is slightly different form that of PI control in that there is no test for Loop Status (6) because, referring to Figure 3-5, there is no one-to-one mapping between R_n and the frequency of the trend. The Loop Status (6) situation is then merged into Loop Status (5), which is termed long-term cyclic. Whilst for PI control, The cyclic is sub-divided into either long-term low-f cyclic or ultimate cyclic based on the value of R_n .

A slightly different situation arises if the set-point is changed whilst data is collected. As explained in Step 1 of Section 3.2.3, the controlled variable data set is normalised with respect to the set-point unless the set-point is changed when mean value μ_y centre-based values are used instead. In this case, the effect of a set-point change can be likened to that of an external disturbance, which would result in a particular Loop Status that reflects the trend style of the set-point change, for example, a step change of the set-point will lead to a *short-term transient* status.

Figure 3-7 shows some results obtained by applying a loop status assessment procedure, which is illustrated in Figure 3-8 as a flow chart. Brief descriptions of the three simple examples are as follows:

- (a) pertains to an oscillating pressure control loop in an oil refinery;
- (b) & (c) pertain to two separate simulations.
 - (b) a slow sensor drift from 1000s;
 - (c) an non-stationary disturbance from 1000s.

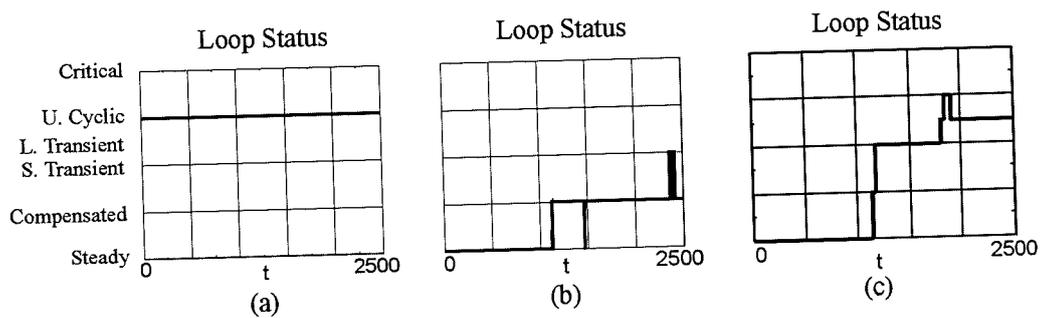


Figure 3-7: Some simple examples of loop statuses

(b) & (c) contain some unexpected '*spikes*', which might mislead or confuse operators when they occur. This leads on to the next section, which examines the possibility of filtering.

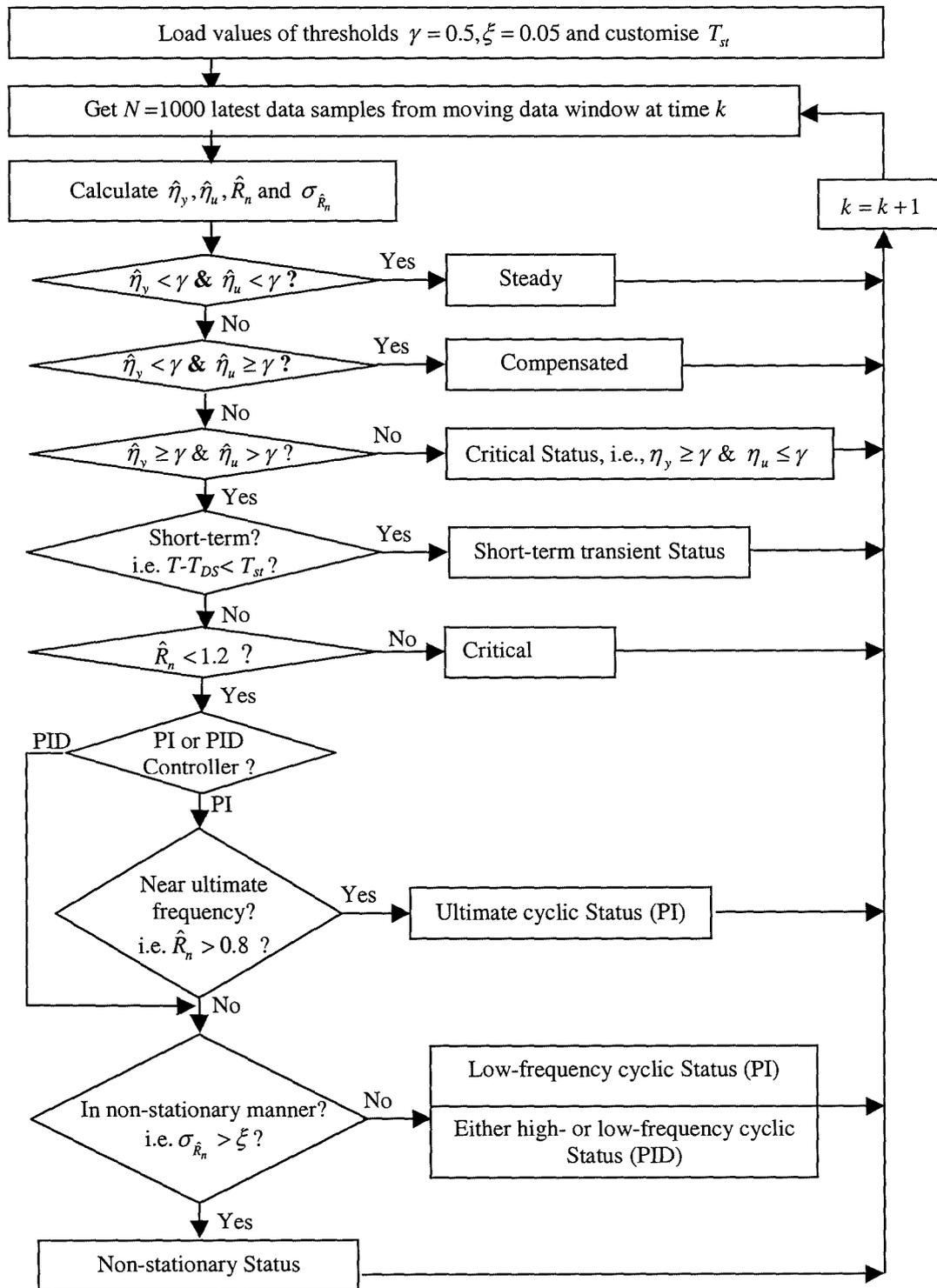


Figure 3-8: Flowchart of loop status assessment

3.3.2 The Quantitative Loop Status Statistic

A quantitative version of Loop Status, the LS statistic, can be found by assigning the scores given in Table 3-4 to each category. The scores are merely chosen to produce a statistic that increases as the loop deteriorates. The quantitative version can then be filtered, for instance by applying a simple EWMA filter:

$$LL(k) = (1 - \lambda)LL(k-1) + \lambda LS(k) \quad (3.21)$$

where $LL(k)$ is the filtered statistic, $LS(k)$ is the score, λ is the filter factor, which is chosen so that it solely suppresses spikes, and k is the time sequence. The benefit of this formulation is that it converges to a value that represents the loop status.

<i>Loop Status</i>	<i>LS</i>
Steady	0
Compensated	1
Short term transient	2
Long term transient	2.5
<ul style="list-style-type: none"> • Cyclic ($\sigma_{\hat{R}_n} \leq \xi_3$) • Non-stationary ($\sigma_{\hat{R}_n} > \xi_3$) 	
Ultimately cyclic	3
Critical	4

Table 3-4: Quantitative values of Loop Status

A direct use of this quantitative statistic would be in the provision of an operator aid. A controller icon on a plant schematic could be made to change its colour or pattern in response to a change in loop status. Thus for instance the icon might be

changed from green to yellow, when the controller has to compensate. Before a controller icon could be displayed properly, the smoothed/filtered LL statistic would have to be converted back to its qualitative description. The conversion rules are given in Table 3-5.

Status	Range
Compensated	$0.5 < LL < 1.5$
Short term transient	$1.5 \leq LL < 2.25$
Long term <ul style="list-style-type: none"> • Cyclic ($\sigma_{\hat{r}_n} \leq \xi_3$) • Non-stationary ($\sigma_{\hat{r}_n} > \xi_3$) 	$2.25 \leq LL \leq 2.75$ (PI) $2.25 \leq LL \leq 3.5$ (PID)
Ultimately cyclic	$2.75 < LL \leq 3.5$ (only for PI)
Critical	> 3.5

Table 3-5: Rules for converting the quantitative statistic into a qualitative version

3.4 Examples

Three examples of what might happen are shown in Figure 3-9: The upper panels show what happened when a slow sensor drift was introduced into a PID loop, the middle panels show the effect of a sticking control valve in a PI loop, and the lower panels show the effect of a load change in a PI loop. The plots show deviations in controlled variables from their set-points, controller output deviations from their initial values and the filtered quantitative statistics LL , all versus time.

A simulated slow sensor drift: the LL value changes to status 1 to indicate that there is some kind of small disturbance or fault, which has been compensated for by the controller.

Control valve stick obtained from a real plant: the LL value is at 2.5 and $\sigma_{\hat{R}_n} = 0.016$, which indicates a low frequency, cyclic trend, which may be caused by a valve problem (sticking) or some external disturbance. Plant operators have confirmed that the valve was sticking.

A simulated change in load: the load change results in a step change in controller output causing the status to increase transiently. This indicates a short-term step-style load disturbance or, perhaps, a sudden change in sensor bias. In practice the controller parameters are no longer optimised.

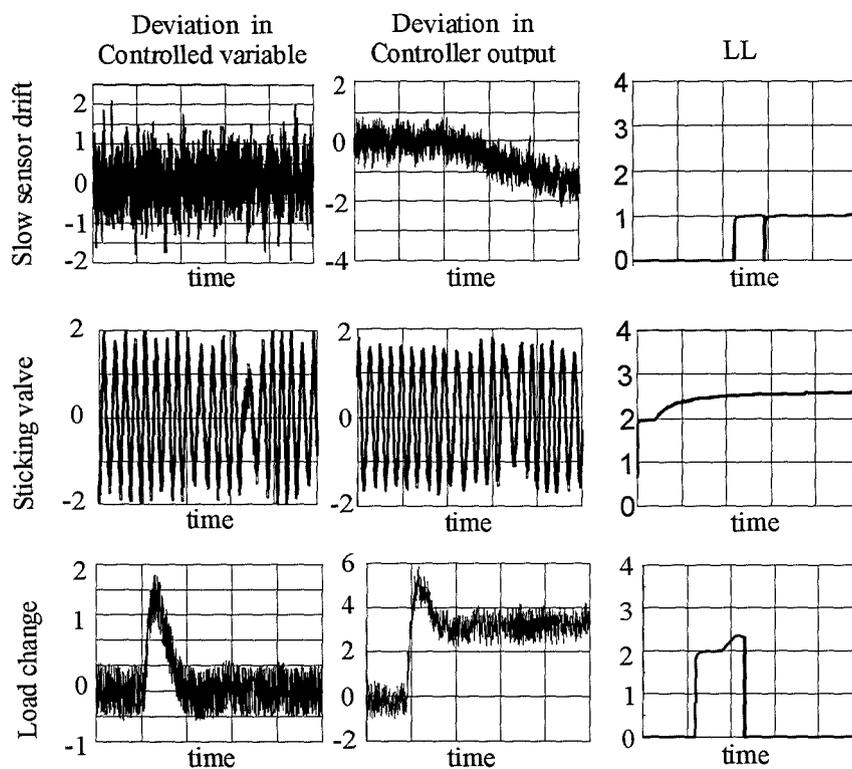


Figure 3-9: Examples of filtered *LL* statistics and icon statuses

Figure 3-10 gives an example of the difference between the long-term sub-categories, non-stationary and cyclic disturbance. Focusing on the R_n plots, there is a more pronounced variation in the non-stationary case, which results in a larger $\sigma_{\hat{R}_n}$. A value of 0.0126 is obtained for the low frequency cyclical disturbance, whilst a value of 0.1158 is obtained when a long-term, non-stationary disturbance is introduced instead. The default ξ of 0.05 would discriminate between these two.

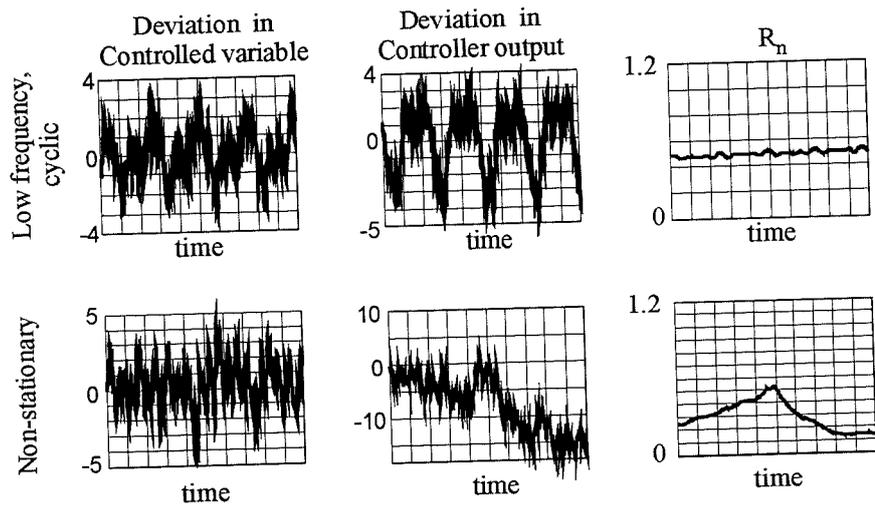


Figure 3-10: Comparison between a low- f oscillation and a non-stationary disturbance

3.5 Summary

This chapter proposes various statistics to facilitate the status monitoring of PI/PID loops. These statistics are based on the signal-to-noise ratio of the controller's input and output data, and on the ratio between them. No process models are needed. Although loop status is a qualitative description, a quantitative version of loop status (*LS*) can be formulated by assigning different real number values to the various categories and a filtered & smoothed statistic (*LL*) by an EWMA filter. It is also envisaged that a controller icon on a plant schematic might be made to change its colour or pattern in response to a change in loop status.

Although the approach has been developed for PI/PID controllers, it should be equally applicable for those controllers that have a frequency dependent statistic *R*, which is the ratio of the signal to noise ratios of the controlled variable and controller output. Thus the approach is not suitable for a P control loop, because *R* will always be unity. Equally it also not feasible for open-loop control, because *R* is intended to reflect the relationship between the controlled variable and the controller output, and this does not exist for an open-loop control system.

The loop status assessment procedure can be embedded as a loop self-validation module into any distributed intelligent PI/PID control nodes, like those in Field-Bus control systems. Or it can be implemented in any high-level supervisory software package, such as a SCADA system.

Loop Status could have a number of different uses:

- as an operator aid; a loop status icon on a plant schematic might provide an overall impression of control activity on a plant;
- as an aid for the control/maintenance engineer: Table 3-1 lists some common causes of Loop Statuses;
- as a pre-processing procedure for the localisation/isolation of the source loop of a fault/disturbance in a unit-wide or plant-wide control system containing a number of interacting loops. The isolation of the problem loop will be considered in the next chapter.

Chapter 4

Fault Localisation in a Number of Interacting Control

Loops

The previous chapter has described how the condition of a PI or PID control loop can be assessed by Loop Status monitoring. The statuses of individual loops cannot be viewed in isolation, however, because in certain circumstances a number of loops might be adversely affected by the degradation of one loop, and this situation might come and go. For instance a plant might operate in a regime where a single loop causes an entire plant to oscillate, but only whilst operating in that particular regime. Both the operator and control/maintenance engineer would be interested in not only locating the problem loop, but also in doing so whilst the situation still exists on the plant. For instance this would enable an engineer to experiment with controller settings.

The quantitative statistic of loop status, i.e., LL and other single number statistics, like η , and η_u , can be combined to form an index to isolate a problem loop in a number of interacting loops. This combination is called the *overall loop performance index (OLPI)*, which increases as loop performance deteriorates. The problem loop can then be isolated by finding the loop with the largest OLPI.

This chapter first describes how these statistics can be obtained from real data. It then goes on to look at how a problem loop can be located in a number of

interacting loops. The method is finally tested on a number of applications, both real and simulated.

4.1 OLPI Derivations & Fault Localisation

The aim is to devise a loop index, the overall loop performance index or *OLPI*, whose value is largest for the problem loop. Thus an index is required that is suitably scaled (for inter loop comparison) and increases with loop degradation. Observe that the larger the value of *LL*, the worse the loop's performance, but this is too coarse a measure for purposes of comparison. On the other hand either η_y or η_u or both will be large for $LS > 0$, so one possibility might be

$$OLPI(k) = \max(\hat{\eta}_y(k), \hat{\eta}_u(k)) / \gamma \quad (4.1)$$

It is intuitively true that the problem loop will suffer the most effect from deterministic disturbance, so the signal-to-noise ratio indexes of the problem loop will be larger than those of affected loops, which are perturbed by some attenuated deterministic disturbances because of the filtering effect of propagation paths. While a more realistic *OLPI* index is that it may also take into account the extent of loop degradation. Thus statistic *LL(k)* is incorporated into Equation (4.1) to reflect loop deterioration:

$$OLPI(k) = LL(k) \max(\hat{\eta}_y(k), \hat{\eta}_u(k)) / \gamma \quad (4.2)$$

Having generated the *OLPI* for each loop, it can be argued, intuitively, that the larger the *OLPI* value of a certain loop, the more likely this loop is to be the source of the problem. This leads to the following localisation procedure:

- check *OLPI* values for each loop within a group of interacting controllers, locate the loop with the largest *OLPI* value, and mark this loop as most probably containing the root cause;
- check the Loop Status of that loop to narrow down the possible causes (refer to Table 3-1);
- if necessary, perform further analyses, such as a Non-linearity Index Test (Thornhill *et al.* 2001) and others (Horch 1999; Thornhill *et al.* 1999), to diagnose the fault or disturbance in that loop, especially for an oscillation problem.

4.2 Application to Real Data From an Eastman Chemical Company Plant

The Eastman Chemical Company plant contains 9 master loops as shown in Table 4-1. Time trends of a 2-day plant oscillation data set are shown in Figure 4-1. It can be seen that most of the plant loops are perturbed by a significant cyclic disturbance as well as some low frequency trend disturbance. *OLPIs* and Loop Statuses of these loops are examined and prove to be successful not only to detect loop oscillation, but also to localise the problem loop.

Num.	Loop Tag	Description of Loop Control
1	PC1	Re-boiler 1 pressure (cascade)
2	LC1	Column 1 level (cascade)
3	FC4	Column 2 feed flow
4	TC1	Column 2 bottom temperature (cascade)
5	PC2	Separator 2 pressure
6	LC3	Column 2 level (cascade)
7	LC2	Decanter 1 level
8	TC2	Column 3 top temperature (cascade)
9	FC7	Column 3 steam injection rate

Table 4-1: Description of 9 master loops in an Eastman Chemical Company plant

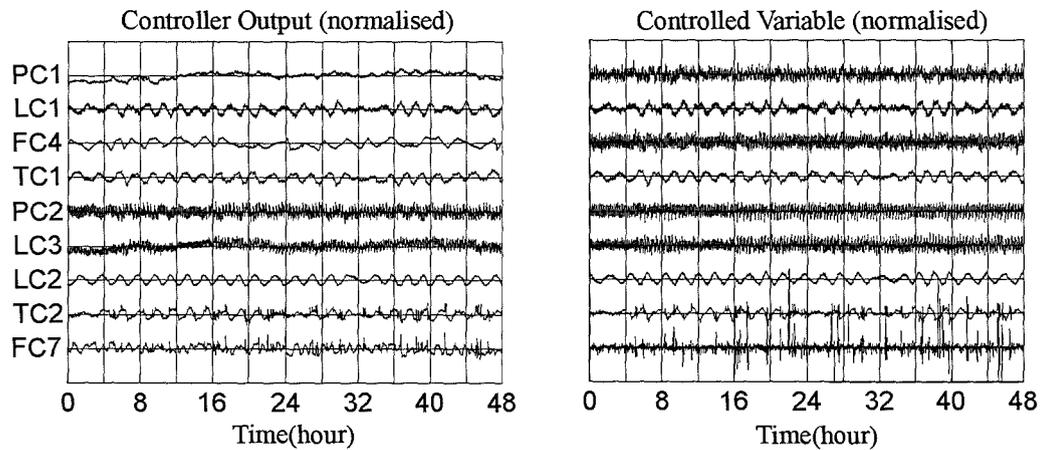


Figure 4-1: Time trends of 9 master loops in an Eastman Chemical Company plant

The Overall Loop Performance Index (*OLPI*) was evaluated for each loop by applying a moving data window of 1000 sample points at each time step assuming that all the set-points remained unchanged during this period. All default thresholds were used. The *OLPI* values were then averaged over the two days. The results are listed in Table 4-2. Detailed analysis results are attached in Appendix B.

The results suggest that loop LC2 contains the root-cause, because of its largest *OLPI* value. Its loop status then indicates that it is operating in a long-term, cyclic mode. A valve problem has been verified by non-linearity test, and has also been confirmed by plant engineers(Thornhill *et al.* 2002b).

Tag	Mean(<i>OLPI</i>)	Loop Status
PC1	11.6	Compensated
LC1	14.3	Long-term cyclic
FC4	48.2	Compensated
TC1	38.4	Ultimate cyclic
PC2	< 5	Ultimate cyclic status detected, temporarily, early on.
LC3	< 5	Compensated
LC2	132	Long-term cyclic
TC2	13	Long-term cyclic
FC7	7	Compensated

Table 4-2: *OLPIs* & Loop Status for an Eastman Chemical Company plant real data set

4.3 Case Studies with the Tennessee Eastman Process Benchmark

The Tennessee Eastman (T-E) process benchmark is a simulation of a real plant that has been disguised for proprietary reasons, the flow diagram and its basic control scheme is shown in Figure 4-2. The process produces two products, G and H, from four reactants, A, C, D and E. Also presents are an inert B and a by-product F. The process has five major units: a reactor, a product condenser, a vapour/liquid separator, a recycle compressor and a product stripper. The gaseous reactants are fed to the reactor where they react to form liquid products. The gas phase reactions are catalysed by a non-volatile catalyst dissolved in the liquid phase. The products leave the reactor as vapours along with un-reacted feeds and the catalyst remains in the reactor. The process has 41 measurements and 12 manipulated variables. A plant-wide decentralised control scheme has been developed by McAvoy *et al* (1994). It is based on multiple single-input-single-output (SISO) control loops. Further details of the process and its control systems can be found in (Downs & Vogel 1993), as well as (Howell *et al.* 1997; McAvoy *et al.* 1995; McAvoy & Ye 1994).

The process plant contains multiple single-input-single-output (SISO) control loops as outlined in Table 4-3: many of these are standard cascade control systems. Table 4-3 also contains values for the threshold parameters γ , two of which were customised by analysing normal steady state data.

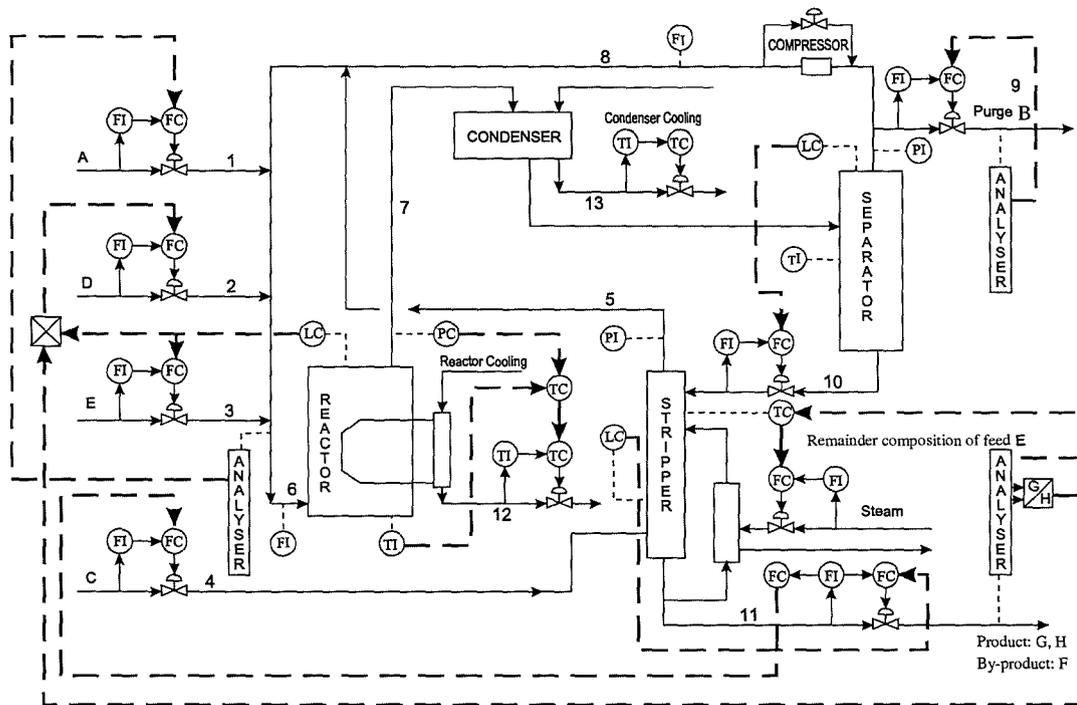


Figure 4-2: Flow diagram and basic control of the T-E process benchmark

Loop	Node	γ
1	A/C composition — A flow rate cascade control	Default
2	G/H composition — reactor level — D,E flow rate control	Default
3	Reactor pressure — temperature cascade control	7*
4	Condenser cooling control	Default
5	Purge composition B — purge flow rate cascade control	Default
6	Separator level — underflow rate cascade control	Default
7	Stripper level — product flow rate cascade control	Default
8	Product flow rate — C flow rate cascade control	Default
9	Product composition E — stripper temperature cascade control	1.4 *

* Loops 3 and 9 have somewhat non-steady initial behaviours, so customised γ values are used.

Note: The loop numbers in this table do not correspond to the numbers appeared in Fig. 4-2 which refer to streams.

Table 4-3: The nine master loops of the TE process benchmark

Disturbances or faults can be injected into the plant. Three common types of disturbance/fault are examined in these case studies: Step, Sensor Drift and Kinetic Slow Drift. More descriptions of possible disturbances or faults can be found in (Chen 2001).

In each case there are two sets of plots: one set shows the original data of the problem loop, whilst the other gives the *OLPI* for the loops of interest, *LL* plots are also given in most cases. In all cases it was difficult to decide whether loops 3 and 9 were badly tuned or were just continually subjected to excessive noisy disturbances.

4.3.1 Case 1: Step Change Disturbance to Loop 5 (B Composition)

The discussion in this section is associated to Figures 4-3 and 4-4. Loop 5 regulates the composition of B by diverting a proportion of the recycle stream out of the process via a purge valve. The perturbed recycle stream then affects the composition into the reactor and this disturbs Loop 1 (a cascade system that regulates the A/C component into the reactor), and the reactor itself. Thus reactor pressure (Loop 3), condenser cooling water temperature (Loop 4) and eventually the product flow rate (Loop 8) are affected. It is easily seen that Loop 5 (Purge composition B— purge flow rate cascade control) has a significantly larger *OLPI* value than others. The *LL* trend of loop 5 indicates that the behaviour is associated with a short-term transient disturbance. After a period of time, Loops 1, 4, 5 all return to good conditions ($LL=0$), whilst Loops 3 and 8 continue to compensate for the change in operating point.

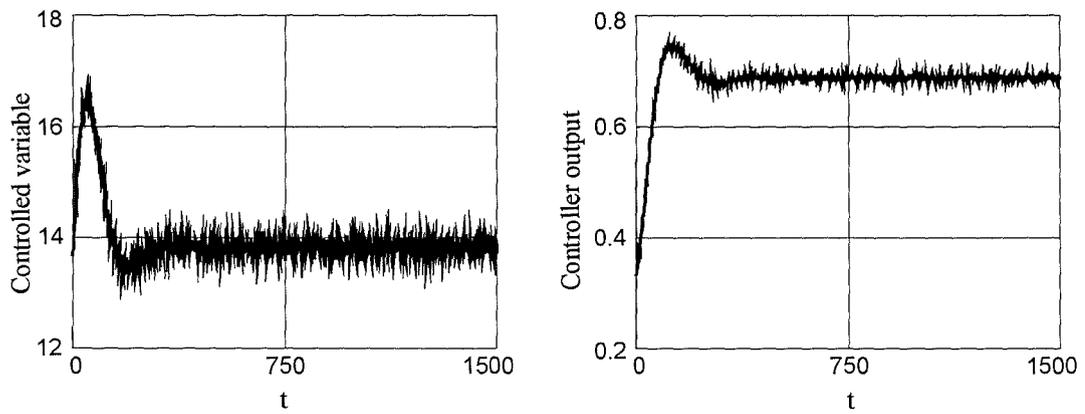


Figure 4-3: Step change disturbance to B composition: Loop 5 data

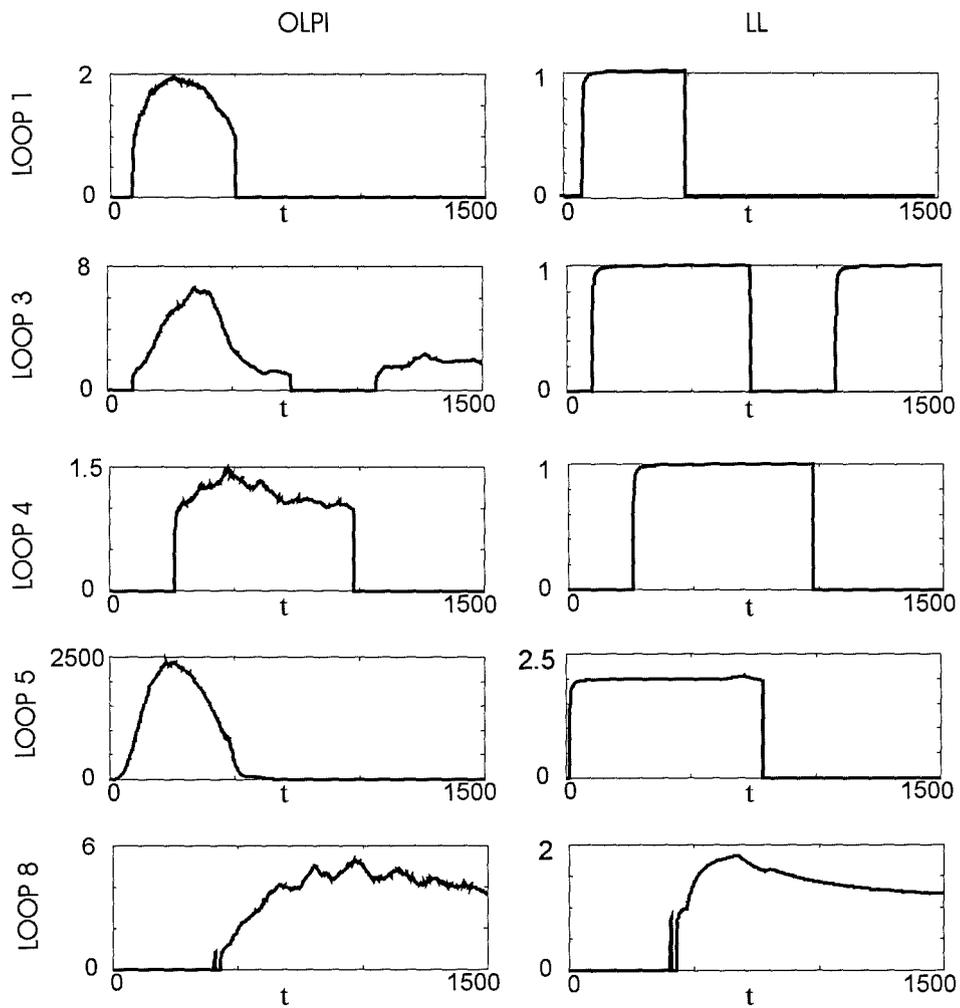


Figure 4-4: Step change disturbance to B composition: *OLPI* & *LL*

4.3.2 Case 2: A-Feed Flow Sensor Drift

The discussion in this section is associated to Figures 4-5 and 4-6. In this case, Loop 1, 5 and 9 are perturbed, while loop 1 has a largest OLPI value. The LL value, which tends to value 1, shows that a small disturbance/fault, such as sensor null shift or slow drift, might have happened in loop 1.

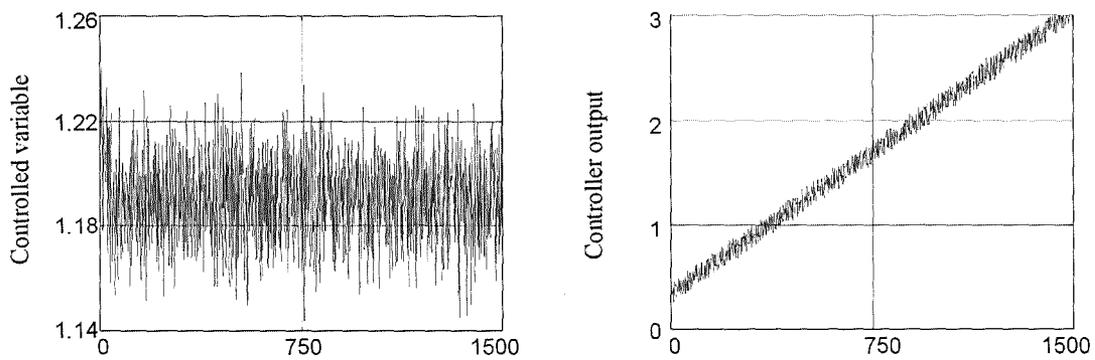


Figure 4-5: A-feed flow sensor drift: Loop 1 data

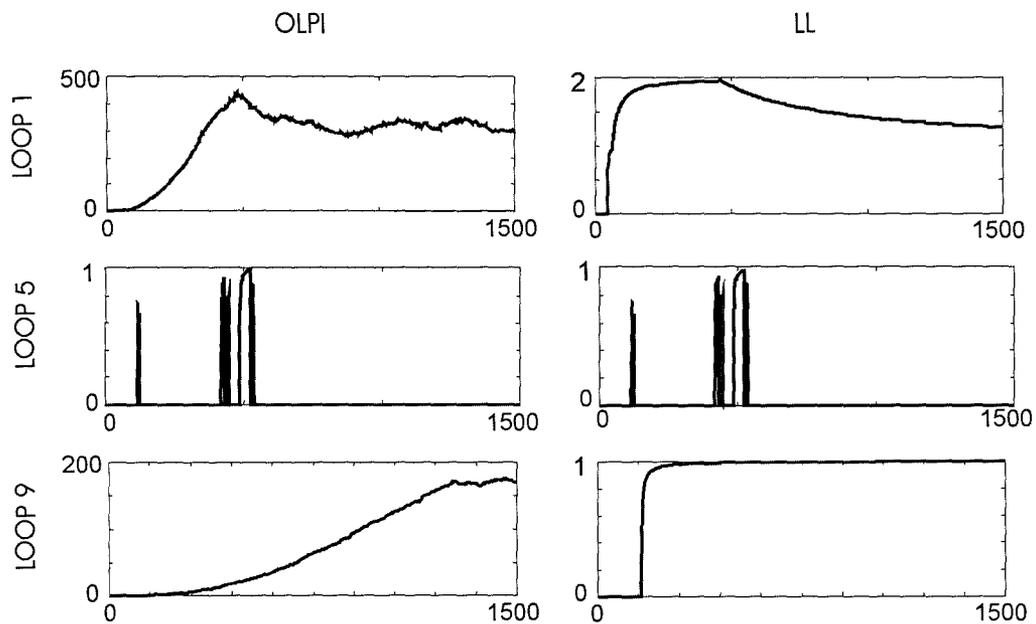


Figure 4-6: A-feed flow sensor drift: OLPI & LL

4.3.3 Case 3: Reaction Kinetic Slow Drift

The discussion in this section is associated to Figures 4-7 and 4-8. The reactor is central in the T-E process plant. Its performance affects the rest of the plant and the plant recycle contributes to the generation of complicated transients. Although Loop 3 (the reactor pressure cascade control loop) is central to its performance, other loops are coupled tightly. Initially the change in reactor kinetics causes a small, slow drift in the OLPI's of at least 5 loops. However the plant then experiences larger, short-term transients and during this period the OPLI of Loop 3 rises, significantly, to about a factor of 10 greater than any other loop. Application of the localisation procedure thus suggests that Loop 3 contains the root cause, because the OLPI is considerably greater than others.

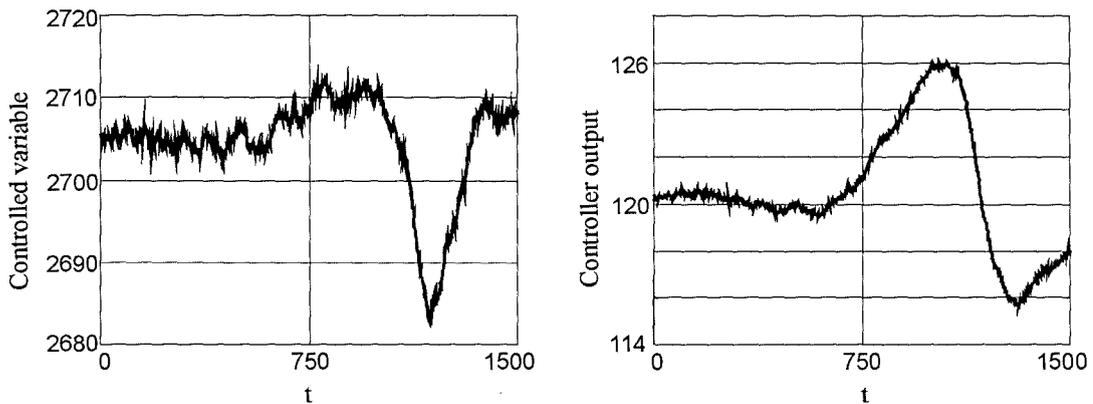


Figure 4-7: Reaction kinetic slow drift: Loop 3 data

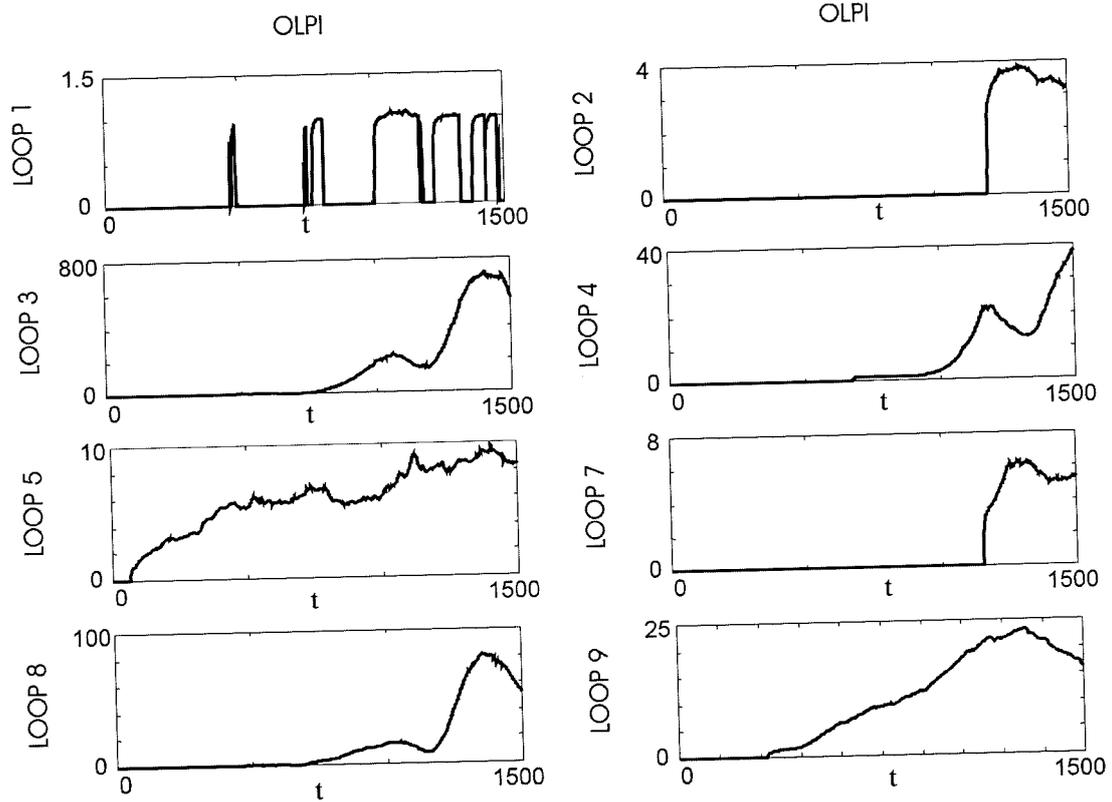


Figure 4-8: Reaction kinetic slow drift: *OLPI*

4.4 Summary

Fault localisation based on *OLPI* comparison has proved to be successful and robust by both simulation and real industrial data analysis. *OLPI* comparison can point out the problem loop in a multi-loop arrangement, and Loop Status information of that problem loop helps to narrow down the possible root cause. This approach is successful when analysing a plant with interacting loops that is perturbed by ONE dominant deterministic disturbance at a time. It is unlikely to work either for multiple faults that result in multiple dominant disturbances, or for poorly designed tight-interacting plants. Chapter 5 will discuss the special case of a plant that is tightly-coupled leading to similar significant *OLPI* values, and an extension to accommodate the multiple oscillations will be demonstrated in Chapter 6.

Chapter 5

The Case When Two Loops Have Similar Significant *OLPI* Values

In the previous chapter, *OLPIs* have been used to locate a problem loop in a number of interacting *PI/PID* control loops by finding that loop with the largest *OLPI* value. Unfortunately there are cases where the *OLPIs* of two loops will be similar and significant:

- there is a single problem, but the two loops are tightly coupled;
- each loop has a separate fault, which coincidentally have similar *OLPI* values.

The first scenario pertains to a tightly coupled plant, whilst the second doesn't. In this chapter, the first scenario will be examined, whilst the second, i.e. multiple faults case, will be studied in the next chapter.

5.1 Tightly Coupled Control Loops

Let the open loop transfer function matrix of the two loops be

$$\mathbf{G}(s) = \begin{bmatrix} g_{11}(s) & g_{12}(s) \\ g_{21}(s) & g_{22}(s) \end{bmatrix} \quad (5.1)$$

In reality this would be operating point dependent so that the individual elements might vary in time. The two loops would then be tightly coupled if both $|g_{12}(j\omega)| \approx |g_{11}(j\omega)|$ and $|g_{21}(j\omega)| \approx |g_{22}(j\omega)|$ over the relevant frequency range.

This situation can arise in process plants, which have recycling or heat recovery systems. Tight coupling might also be operating point dependent so that sometimes plant behaviour might be acceptable, and at other times unacceptable. These situations are largely attributable to poor plant design, and the solution might well involve re-design.

One of the symptoms of tightly coupled loops is that they behave as one. This is of particular importance if the plant is experiencing, what might be loosely termed, a 'whole plant oscillation'. Whole plant oscillations are caused in at least three different ways:

- a single problem loop acts as an oscillator or resonator affecting the entire plant; Loop Status monitoring and OLPI-based localisation reasoning, as discussed in Chapter 4, can be used to detect and localise the problem loop. Thornhill *et al.* (2002b) have also examined this problem by combining data-driven techniques, such as non-linearity signature and spectral power test, and process understanding;
- a sinusoidal external disturbance affects the entire plant; this will result in a similar situation to the above: OLPI-based localisation reasoning is also feasible for this situation;
- two or more, tightly coupled control loops act as an oscillator or resonator. Only this third way will result in similar OLPI values. Unfortunately detailed knowledge of the plant would be needed to explain the physical phenomena that causes the oscillation or resonance. The analysis here merely associates the problem with poor plant design.

An example: a cascaded, two CSTR system can be made to exhibit coupled behaviour if a heating/cooling water heat recover system is connected between them (Figure 5-1). Each CSTR has both temperature and level control. The up-stream CSTR carries out an endothermic chemical reaction, whilst the down-stream CSTR, is an exothermic one.

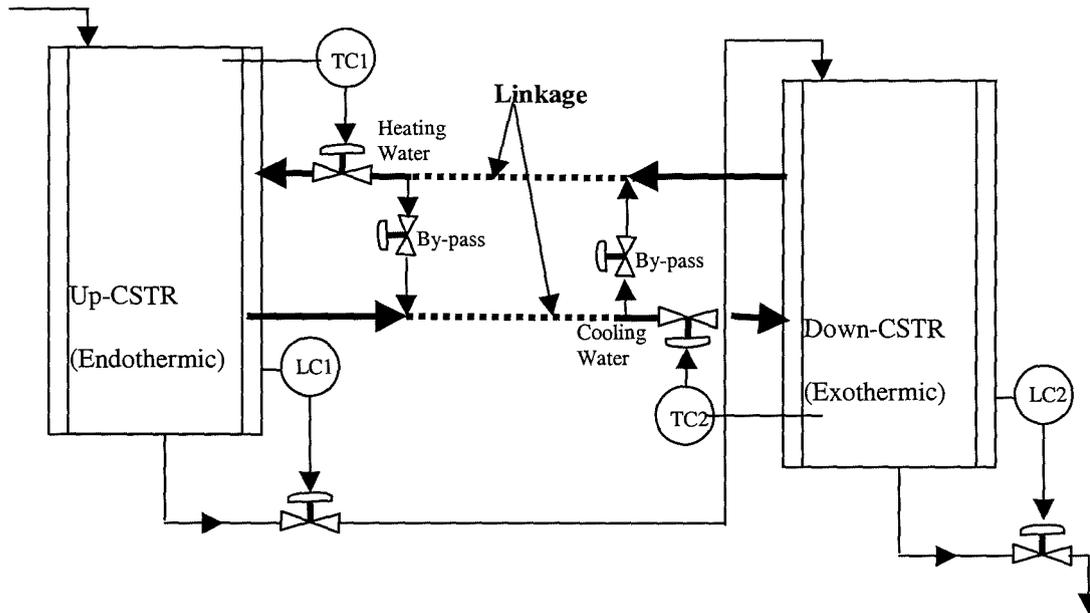


Figure 5-1: Plant schematic of two cascaded CSTRs

It is common practice to install Heating/Cooling systems, for the purpose of energy conservation. In the example here it has been deliberately contrived so that both CSTRs are designed without any consideration of heating/cooling recycle linkage. When this system is commissioned, both TC1 and TC2 oscillate. Detailed model information and figures of operating data can be found in Appendix C.

5.2 Commissioning-Stage Poor Performance (CSPP)

Both model-based and data-driven fault detection & diagnosis techniques usually have an implied assumption that the plant/system monitored will be in a normal, steady and healthy condition, prior to the occurrence of a change or fault. This is normally true because most of the plants work well after proper commissioning. But about 15% of plant control loops are found to have poor performance during commissioning. They may exhibit slightly under-damped behaviour, excess variability or even oscillate. This may be due to poor plant or control. (Note that the commissioning-stage here includes periods when there are significant changes to the operating points). A consequence of this is that a plant might need to be redesigned. A possible reason of this Poor Design Problem (PDP) could be the lack of consideration of a strong cooling/heating recycle linkage between sub-systems within a plant.

5.2.1 Poor Design Vs. a Commissioning-Stage Fault

Consider the situation of the detection of poor performance during commissioning. It would be relatively easy to diagnose the problem if it is attributable to one loop, whilst harder to get a satisfactory answer when many interacting loops are involved. There is therefore a need to:

- (1) prevent designers from doing massive redesigning work if it's just a commissioning stage fault;
- (2) give some guiding information about what's wrong with the process structure and/or control strategy;

- (3) help designers to concentrate on key structure problems in their redesign work.

Model-based and PCA-related detection methods are not suitable for this problem because the 'normal' model, either an analytical or a statistical model, generated from normal operation data, isn't available during this period.

This chapter seeks to find a way of discriminating between poor plant design and a commissioning stage fault in a plant containing a group of conventional PI/PID control loops. The investigation is data-driven so only sensor measurements are needed. Detailed model information would be of only limited usefulness because poor plant design implies that something has not been taken into account in the modelling.

5.2.2 Test of Control Efforts

When multiple interacting conventional PI/PID control loops are involved, each controller will attempt to accommodate for the excessive variability. An index quantifying the '*level*' of control effort of each control loop could be useful to help to isolate the cause. The *OLPI* gives a proper quantitative measure of control effort and grows larger as the problem worsens. Alternatively, the signal-to-noise ratio of controller output can be used as a control effort index, when non-PI/PID loops are involved.

Application of the principle of pairing of controlled and manipulated variables (Stephanopoulos 1984) would result in a controller which reacts more strongly to a

change/fault which happened within its loop than within another interacting loop. It is because the relative gain between the controller output of a faulty loop and the controlled variable of any other loop should be smaller than the direct gain in the faulty loop. Thus those interacted loops need only give relatively less efforts to deal with the affect. This is normally true unless ill-conditioned pairings or tightly coupled interactions exist. In such special cases, several loops could show similar control efforts, i.e., *OLPI* values. An intuitive conclusion can be made as follows:

- (1) if a loop has '*much larger*' *OLPI* value than others, it could either be due to a fault or a bad-tuning in that loop;
- (2) if two or more loops have '*similar*' *OLPI* values, the local problem area could be as a result of a design problem, such as hidden tightly coupled interactions.

For conclusion (1), a further test is needed to distinguish between whether the loop is suffering from a bad-tuning problem or a fault, such as a valve or sensor problem. The Loop Status statistic and/or oscillation diagnosis analysis (Thornhill *et al.* 2001) can be used for this purpose.

Terms '*similar*' and '*much larger*' can not be defined generally, because the ratio of *OLPI* values has a close relationship with the dynamic interactions, which could be different for different situations. An attempt was made to identify the dynamic interaction by just analysing closed-loop running data, and then finding its affect on the value of *OLPI*. This failed and it was concluded that specific tests would be needed like the application of dithering signal, or an open loop step test. One can

only get some interaction information from design-stage specifications. If it is known that interactions between loops are not extremely strong, a threshold value of 2, can be used. That is to say, if the ratio of the largest two OLPI values is smaller than 2, the two OLPI values could be thought of as '*similar*', otherwise one is '*much larger*' than the other.

5.2.3 Test of Uniqueness of Dominant Spectral Components

If a particular part of the plant is suspected of containing the source of plant-wide oscillations caused by poor plant design, the spectral content of any of the signals recorded from this area are likely to contain the same spectral dominant components. Multiple disturbances are likely to result in multiple frequency components. The same location of dominant peaks in these spectra implies the uniqueness of the dominant component, otherwise multiple faults/disturbances could occur in the system. A spectral PCA based clustering method (Thornhill *et al.* 2002a) can be used to see if these loops are suffering from the same dominant component, when many loops contain similar OLPI values. Otherwise, a comparison of the peak values of power spectra is normally sufficient.

5.2.4 Test of Physical Properties

Although tests of control effort and uniqueness of the dominant spectral component may help commissioning engineers to have an idea of whether there is a design problem within a physical sub-system area, they cannot provide any insights into a design that might be flawed. A greater physical understanding is required to

progress further, for instance of energy flows or mass flows. A test is proposed here to help to determine whether there is an intrinsic physical linkage, such as a strong cooling/heating recycle, which can cause the poor performance in the problem area. The procedure for such a test is plant specific, but could have the following guiding steps:

Step 1: Locate the problem physical area, i.e. the area that contains those poorly performing loops with '*similar*' OLPI values and the same dominant spectral components, then determine the physical property of interest, i.e. energy or mass trend based on the criteria in Table 5-1.

Loops involved	Type of test
Temperature	Energy trend
Level , Flow	Mass trend

Table 5-1: Criteria of physical test

Step 2: Divide the problem area into several stand-alone units, and, based on available measurements, calculate the variation in the net energy or mass stored in each unit by simply considering each unit as a black box.

Step 3: Check the cross-correlation of these energy (mass) trends to see whether there are strong linkages between them, and verify whether these linkages have been fully reflected in the design model.

Step 4: Give some advice of how to improve the performance by redesigning either control strategy or structure or both.

5.3 Case Studies

Two cases are used here to illustrate the tests on the commissioning-stage plant data obtained from a simulation of the two cascaded CSTRs discussed above (see Appendix C for more details). Case 1 shows how the tests introduced in the previous section can point to a process design problem and its root cause, while Case 2 demonstrates a fault.

5.3.1 Case1: Tightly Coupled Through Recycling

The two cascaded CSTRs are commissioned using pre-specified control parameters without any consideration of the effect of the heating/cooling recycle. It is found that both TC1 and TC2 suffer from oscillations, which are not expected to happen during at design stage.

OLPI Test (Control efforts)

Table 5-2 shows the estimated OLPI values for the TC1 & TC2 loops from oscillatory operating data. Similar OLPI values indicate that a design problem may exist in the system.

<i>Loop</i>	<i>OLPI</i>
TC1	27
TC2	36

Table 5-2: OLPI Values for TC1 &TC2

Power Spectra Test of TC1 & TC2 Controller Outputs

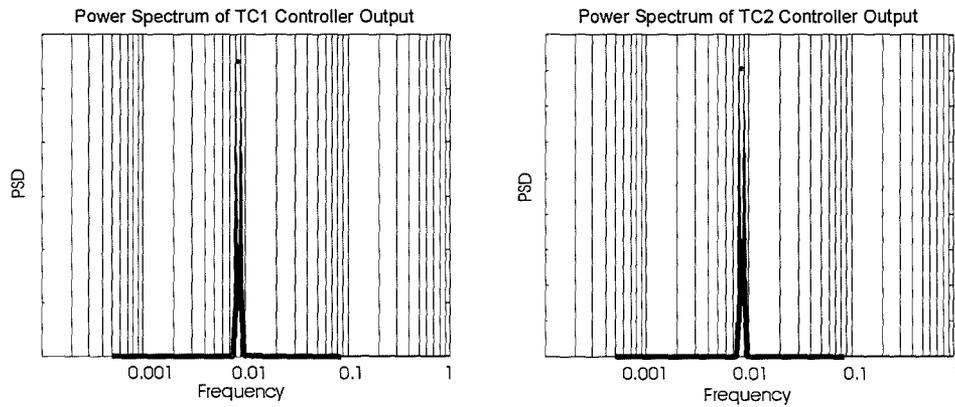


Figure 5-2: The power spectra of TC1 & TC2 controller outputs

It is quite clear from the power spectra of TC1 & TC2 controller outputs (Figure 5-2) that they are suffering from the same dominant component due to the same location of their peaks. The uniqueness of the dominant component implies a process design problem rather than multiple-faults/disturbances.

Transient Energy Trend (Power Trend) Test

Because the problem is associated with temperature, an energy trend test is of interest. Each CSTR can be thought of as a stand-alone unit, and net power trends for the two CSTRs can be calculated using heating/cooling water flow and temperature measurements, as well as inlet reactant and outlet product flow/temperature measurements. Figure 5-3 shows power trends and their cross-correlation.

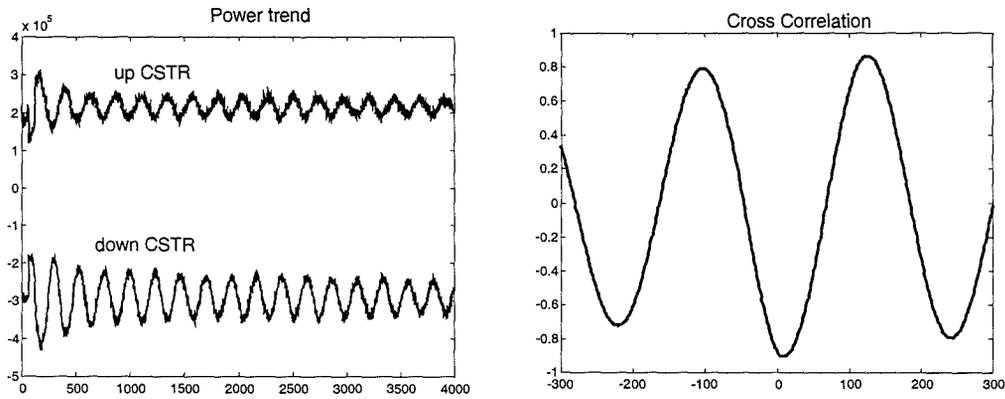


Figure 5-3: Power trends and cross correlation

It is clearly seen that the two CSTRs have oscillatory power trends with the same frequency, and also have a strong negative relation at lag zero, which means energy in the system is switched between the up and down CSTRs. According to the plant schematic, a conclusion can be made that the strong and direct heating/cooling recycle linkage causes the oscillatory behaviour in the system. The solution is to try to find a way to weaken this linkage. One successful way is to put a buffer in the heating/cooling linkage, so as to reduce the variability of the heating/cooling inlet recycle temperature of each CSTR.

5.3.2 Case 2: Valve Problem in LC1

In this case, the tightly coupled problem in Case 1 has been solved by putting a water buffer in the recycle stream to break the strong coupling. A sticking valve is then simulated and the test result is as follows:

<i>Loop</i>	<i>OLPI</i>
LC1	250
LC2	100

Table 5-3: OLPI values for LC1 & LC2 with valve problem in LC1 loop

The OLPI value of LC1 is *much larger* than that of LC2. So it is most likely a fault problem. The further test according to valve non-linearity is done, and a non-linear valve problem has been verified.

5.4 Conclusions

A less emphasised problem, Commissioning-Stage Poor Performance (**CSPP**) caused by tightly-coupled loops, has been discussed here. A data-driven diagnosis method for CSPP problem is proposed. Although fairly simple, case study tests show its capability to distinguish between a poor design problem and a commissioning-stage fault, and to reason about poor process design. It is worth pointing out that a combination of a poor design problem and commissioning-stage fault may cause a wrong conclusion. Verified plant knowledge could also be helpful yet is not considered here.

Chapter 6

Detecting and Isolating Multiple Plant-Wide Oscillations

A new application of independent component analysis (ICA), called spectral ICA, is proposed in this chapter to detect and isolate the sources of multiple oscillations in a chemical process plant. Its key feature is that it extracts dominant spectrum-like independent components each of which has a narrow-band peak that captures the behaviour of one of the oscillation sources. The outcome is a method for determining the origin of each oscillatory disturbance present in the plant and for distinguishing between the sources and their propagated secondary effects that may be observed elsewhere in the plant. Additionally, a significance index is presented that links the sources to specific plant measurements in order to facilitate the isolation of the sources of the oscillations. A case study is presented that demonstrates the ability of spectral ICA to detect and isolate multiple dominant oscillations in different frequency ranges in a large data set from an industrial chemical process. The success of the study shows that spectral ICA represents a useful advance and that the additional computational steps involved compared to a previously presented method for multivariate spectral analysis give benefits in enhanced isolation of the oscillatory sources.

6.1 Introduction

Independent Component Analysis (ICA) is a statistical and computational technique for revealing hidden factors or modes of behaviour within a data set. It can extract independent sources given only observed data that are mixtures of the unknown sources. What distinguishes ICA from other methods is that it looks for components that are both *statistically independent* and *non-Gaussian*. The ICA technique is a relatively new invention. It was introduced in the early 1980s, and several research groups presented algorithms in the mid-1990s, which were demonstrated on problems like the *cocktail-party* effect, in which the challenge is to isolate independent voices in a crowd (Hyvärinen *et al.* 2001). Fields of application included signal processing, advanced statistics and neural computing. Reported applications of ICA include biochemical signal processing, audio signal separation, telecommunications and feature extraction (Comon 1994; Delfosse & Loubaton 1995; Hyvärinen *et al.* 2001; Hyvärinen & Oja 2000; Jutten & Herault 1991; Lee 1998; Sanchez A 2002). Li & Wang (2002) gave the first example of an application of ICA to the analysis of dynamical chemical process trends and dimension reduction. Kano *et al.* (2002a) have shown that time-domain multivariate statistical process control based upon ICA has superior fault detection performance over other SPC methods for a four-variable problem. ICA is often applied to the discovery of unknown sources whose outputs are time series (e.g. Hyvärinen and Oja 2000). However the time-domain approach requires special treatments, such as time-shifting, when the time series have phase lags or time delays that cause the misalignment of similar features. A contribution of this chapter is that these disadvantages can be overcome if ICA is applied to the power spectra of the time trends instead. Additionally, it is proven that spectra having distinct spectral peaks,

for instance from multiple independent oscillatory sources, are maximally non-Gaussian. Thus an ICA technique based on the maximisation of non-Gaussianity is applicable to the discovery of the sources of multiple oscillations.

Section 6.2 presents the methods of spectral ICA and also describes a new technique for multi-range spectral ICA that aids the detection of high frequency oscillations of small magnitude. New insights are also given into why spectral ICA is able to find independent components that extract and isolate individual narrow-band spectral peaks in a data set that is characterised by multiple oscillations. Scaling of the independent components (ICs) is discussed and two new indexes are defined. One of the indexes matches the source of an oscillation to a specific location in the plant and the other gives an indication of the seriousness of the oscillation in terms of its percentage power.

The spectral ICA algorithm is initialised using the results of a spectral principal component analysis. Spectral ICA is applied in section 6.3.1 to a worked example of simulated coupled control loops in which two oscillatory sources are successfully detected and isolated. The results of spectral PCA are compared to the enhanced results from spectral ICA in order to illustrate the added benefits provided by spectral ICA. Section 6.3.2 applies spectral ICA to a large industrial data set in which several independent oscillations are detected and their likely sources are isolated. The analysis involved not only the plant measurements but also the controller output signals that drive the plant actuators, an approach which is shown to give additional insights.

6.2 Methods

6.2.1 Time Domain ICA

Descriptions of time domain ICA models can be found in the literature (for example, Hyvärinen *et al.* 2001; Lee 1998). The Instantaneous Mixing Model (IMM) is one of the more basic models, because it neglects any time delays that might be present in the mixing. This model has the following structure:

$$\mathbf{X} = \underline{\mathbf{A}}\mathbf{S} \quad (6.1)$$

where matrix $\mathbf{X}=[\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_m]'$ represents all sensor observations, $\mathbf{S}=[\mathbf{s}_1 \mathbf{s}_2 \dots \mathbf{s}_n]'$ is the matrix containing all independent, non-Gaussian source outputs and $\underline{\mathbf{A}}$ is the mixing matrix formed with elements $a_{ij}(i=1\dots m, j=1\dots n)$. Matrix \mathbf{A} is underlined to distinguish it from the scaled spectral mixing matrix that will be referred to at the end of section 6.2.2 and used subsequently. Each row of \mathbf{X} , i.e., $\mathbf{x}'_i(i=1\dots m)$, is a vector of observations recorded by the i^{th} sensor for a period of time, and the j^{th} row in \mathbf{S} , i.e., $\mathbf{s}'_j(j=1\dots n)$, is a vector representing the j^{th} independent source output during that period of time, and is termed as an *independent component* (IC). It is clear that each sensor observation is decomposed into a linear combination of a set of independent components, i.e.,

$$\mathbf{x}'_i = a_{i1}\mathbf{s}'_1 + a_{i2}\mathbf{s}'_2 + \dots + a_{in}\mathbf{s}'_n, \quad i = 1\dots m \quad (6.2)$$

In the cocktail party problem the sources are the n individual speakers in the room while the sensor observations are the sounds recorded by m microphones in the room. In a process application, the sources are oscillatory disturbances such as control loops in a limit cycle and the sensors are the plant instruments and other monitored signals. Independent component analysis involves the estimation of both the mixing matrix, $\underline{\mathbf{A}}$ and the independent source matrix, $\underline{\mathbf{S}}$.

In order to simplify the estimation, sensor observations $\underline{\mathbf{X}}=[\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_m]'$ are first centred, i.e., subtracted from their mean vector $\underline{\boldsymbol{\mu}}=[\mu_1, \mu_2 \dots \mu_m]'$ and thus $E(\mathbf{x}_i)=0$ ($i=1 \dots m$). Then the centred $\underline{\mathbf{X}}$ is pre-whitened by a proper linear transformation: $\tilde{\underline{\mathbf{X}}}=\underline{\mathbf{U}}\underline{\mathbf{X}}$, where $\underline{\mathbf{U}}$ is a sphering matrix, such that pre-whitened observations in $\tilde{\underline{\mathbf{X}}}=[\tilde{\mathbf{x}}_1 \tilde{\mathbf{x}}_2 \dots \tilde{\mathbf{x}}_m]'$ are orthogonal, i.e. $\tilde{\mathbf{x}}_i$ ($i=1 \dots m$) are mutually uncorelated and each has unit variance. The sample covariance of $\tilde{\underline{\mathbf{X}}}$ is an identity matrix $\underline{\mathbf{I}}$: $\frac{1}{N-1} \tilde{\underline{\mathbf{X}}}\tilde{\underline{\mathbf{X}}}'=\underline{\mathbf{I}}$, where N is the number of observations, i.e. the column number of $\tilde{\underline{\mathbf{X}}}$. This pre-whitening transformation can be achieved by principal component analysis (PCA) using singular value decomposition (SVD) or eigenvalue decomposition (EVD). The ICA algorithm is normally applied to the pre-whitened observations $\tilde{\underline{\mathbf{X}}}$ first, i.e.,

$$\tilde{\underline{\mathbf{X}}}=\tilde{\underline{\mathbf{A}}}\underline{\mathbf{S}} \quad (6.3)$$

then it is converted back to original mixing model: $\underline{\mathbf{X}}=\underline{\mathbf{U}}^{-1}\tilde{\underline{\mathbf{X}}}=\underbrace{\underline{\mathbf{U}}^{-1}\tilde{\underline{\mathbf{A}}}}_{\underline{\mathbf{A}}}\underline{\mathbf{S}}=\underline{\mathbf{A}}\underline{\mathbf{S}}$. Note that it is based on mean centred values. To obtain the true source vector, $\underline{\mathbf{S}}$ should be corrected by adding $\underline{\mathbf{A}}^{-1}\underline{\boldsymbol{\mu}}$.

The goal of ICA is equivalent to finding a separating matrix \mathbf{W} that satisfies

$$\hat{\mathbf{S}} = \mathbf{W}\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{w}'_1 \\ \mathbf{w}'_2 \\ \vdots \\ \mathbf{w}'_n \end{bmatrix} \tilde{\mathbf{X}} \quad (6.4)$$

where $\hat{\mathbf{S}} = [\hat{\mathbf{s}}_1 \hat{\mathbf{s}}_2 \dots \hat{\mathbf{s}}_n]'$ is the estimate of independent source \mathbf{S} , and \mathbf{w}_j ($j = 1 \dots n$) are separating vectors. Note that \mathbf{W} is orthonormal if all independent sources are bounded to be unit sample variance ($\frac{1}{N-1} \hat{\mathbf{S}}\hat{\mathbf{S}}' = \mathbf{I}$). This is because:

$$\hat{\mathbf{S}}\hat{\mathbf{S}}' = (\mathbf{W}\tilde{\mathbf{X}})(\mathbf{W}\tilde{\mathbf{X}})' = \mathbf{W}\tilde{\mathbf{X}}\tilde{\mathbf{X}}'\mathbf{W}' = (N-1)\mathbf{W}\mathbf{W}' \quad (6.5)$$

then

$$\mathbf{W}\mathbf{W}' = \frac{1}{N-1} \hat{\mathbf{S}}\hat{\mathbf{S}}' = \mathbf{I} \quad (6.6)$$

Thus $\mathbf{w}'_j \mathbf{w}_j = 1$ ($j = 1 \dots n$) and $\mathbf{w}'_j \mathbf{w}_k = 0$ ($j \neq k$). Note that the orthonormal constraint applied on the separating vectors \mathbf{w}_j ($j = 1 \dots n$) in the ICA algorithm guarantees that each estimated $\hat{\mathbf{s}}'_j$ has a unit variance, i.e. it is normalised.

The estimation problem is reduced to finding a set of orthonormal separating vectors \mathbf{w}_j ($j = 1 \dots n$) through pre-whitening. One way of achieving this is to estimate the first separating vector \mathbf{w}_1 , which to maximise the non-Gaussianity of $\hat{\mathbf{s}}'_1 = \mathbf{w}'_1 \tilde{\mathbf{X}}$, and then to estimate the second separating vector on the same basis, and so on. To prevent different vectors \mathbf{w}_j ($j = 1 \dots n$) from converging to the same maxima, $\hat{\mathbf{s}}'_j$ should be de-correlated with those sources $\hat{\mathbf{s}}'_k$ ($k = 1 \dots j-1$) that have been estimated already after every iteration (Hyvärinen and Oja 2000).

The idea that non-Gaussian sources can be extracted by the maximisation of non-Gaussianity, can be explained by the Central Limit Theorem. Any variable composed of several, independent, non-Gaussian random variables will be closer to Gaussian than any single independent random variable. From Equations(6.3) and (6.4):

$$\hat{\mathbf{s}}'_j = \mathbf{w}'_j \tilde{\mathbf{A}} \mathbf{S} = \mathbf{z}'_j \mathbf{S} \quad (6.7)$$

where $\mathbf{z}'_j = \mathbf{w}'_j \tilde{\mathbf{A}}$. Thus the estimated source $\hat{\mathbf{s}}'_j$ is a linear combination of the real hidden independent sources $\mathbf{S}=[\mathbf{s}_1 \mathbf{s}_2 \dots \mathbf{s}_n]'$, and it will be more Gaussian than any of the $\mathbf{s}'_j (j=1 \dots n)$ unless it equals one of the \mathbf{s}'_j .

Non-Gaussianity is measured by either the kurtosis of $\hat{\mathbf{s}}'_j$, (see Equation (6.10)) or by another index like negentropy. ICA usually focuses on time-domain data sets. The kurtosis of a time series whose samples have a Gaussian probability density function is zero, and is non-zero for other time series. By maximising kurtosis, the method seeks to extract dominant non-Gaussian signals.

6.2.2 Basic Spectral ICA

Introduction

The ICA of power spectra, derived from time series, has not been considered before. An example of a dominant non-Gaussian signal in this context might be a spectrum with a single dominant peak. In spectral ICA, different ICs usually contain different peaks because, and as will be shown later, a component having a single spectral peak is more dominant in terms of non-Gaussianity than a component having multiple peaks. The maximisation of non-Gaussianity in spectral

ICA results in the preferential extraction of single-peak, narrow-band independent components rather than multiple-peak, or wide-band components. Thus a spectrum can be decomposed into a combination of spectrum-like and single-peak ICs by means of spectral ICA.

The extraction of dominant narrow-band peaks from a set of power spectra is superior to the extraction of oscillatory sources by time domain ICA, because spectral ICA is invariant to time delays and phase lags, which is not the case with time-domain ICA. For example $\sin(\omega t)$ and $\cos(\omega t)$ will have the same spectrum and will be treated as one source, whilst they are treated as two different sources in time domain analysis. Spectral ICA can be thought of as an extension or improvement of spectral principal component analysis (spectral PCA) as proposed by Thornhill *et al.* (2002a). The disadvantage of spectral PCA is that more than one peak may appear in the spectrum-like principal components (PCs) and different PCs can contain the same peaks, or peaks at the same position but of different sign. Although PCs are orthogonal, i.e. uncorrelated, they are not independent.

In the spectral ICA model, the rows of the data matrix, \mathbf{X} , are now single-sided power spectra $P(f)$ of the observations over a range of frequencies up to the Nyquist frequency (one-half of the sampling frequency):

$$\mathbf{X} = \begin{matrix} \text{N frequency channels} \rightarrow \\ \begin{pmatrix} P_1(f_1) & \cdots & P_1(f_N) \\ \cdots & \cdots & \cdots \\ P_m(f_1) & \cdots & P_m(f_N) \end{pmatrix} \\ \begin{matrix} m \text{ process} \\ \text{variables} \\ \downarrow \end{matrix} \end{matrix}$$

The power spectrum of a signal can be estimated by applying fast Fourier transform (FFT) methods, or the averaged periodogram method (Welch 1967) to the mean centred time trends. The number of frequency channels, N , *i.e.* the row length of \mathbf{X} , is determined by the number of FFT data points (\tilde{N}): $N = \frac{\tilde{N}}{2}$ where \tilde{N} is even in this formulation. Each power spectrum is normalised, *i.e.*,

$$\sum_{k=1}^N P_i(f_k) = 1, \quad i = 1 \dots m \quad (6.8)$$

Formulation

If there are n dominant oscillatory sources, which are independent of each other, then each source will result in a narrow-band frequency peak in one or more of these spectra. Each power spectrum signature will then be made up of a number of these components plus some other residual components that result from normal operation and disturbances. The task is then to extract these narrow band independent components by analysing the spectral signatures using spectral ICA. The narrow band ICs obtained in this way are not unique in terms of sign and magnitude.

To achieve this, \mathbf{X} is first decomposed into n dominant independent components:

$$\mathbf{X} = \underline{\mathbf{A}}_{m \times n} \begin{bmatrix} \mathbf{y}'_1 \\ \mathbf{y}'_2 \\ \vdots \\ \mathbf{y}'_n \end{bmatrix} = \underline{\mathbf{A}}_{m \times n} \mathbf{Y} \quad (6.9)$$

where \mathbf{y}'_j ($j=1 \dots n$) is the j^{th} IC, $\underline{\mathbf{A}}$ is an $m \times n$ real mixing matrix and $m \geq n$.

Following Hyvärinen and Oja (2000), independent components can be estimated by finding a set of orthonormal separating vectors, which maximise the kurtosis of

each normalised independent component in turn. The normalised kurtosis for an centred independent component y is given by:

$$kurt(y) = \frac{\frac{1}{N} \sum_{k=1}^N y^4 - 3 \left(\frac{1}{N} \sum_{k=1}^N y^2 \right)^2}{\left(\frac{1}{N} \sum_{k=1}^N y^2 \right)^2} \quad (6.10)$$

where N is the data length. Computer programs that perform ICA, e.g. FastICA packages, are available in the public domain (Hurri *et al.* 1998).

Insight

To explain why the maximisation of kurtosis can extract a narrow band IC from spectra data, the kurtosis of the power spectrum of Gaussian noise may be compared with either the power spectrum of a pure sine wave or the power spectrum of a compound sinusoidal signal composed of a fundamental and several harmonics. It is shown in the Appendix D that the power spectrum of a finite sample of Gaussian noise has a kurtosis of 6, whereas the power spectrum of a pure sinusoid has a kurtosis of $\frac{\tilde{N}}{2} - 3$ and the power spectrum of the compound signal

has a kurtosis that ranges from $\frac{\tilde{N}}{2k} - 3$ to $\frac{\tilde{N}}{2} - 3$ where k is the total number of sinusoidal components. If \tilde{N} is chosen large enough (>1000), the kurtosis of a spectrum with narrow-band peaks is found to be significantly greater than 6, especially when a hidden source is generating a relatively pure sinusoid signal.

Post-ICA Algorithm

The ICs obtained from the ICA formulation above are not unique in terms of sign and magnitude, so additional constraints are needed to arrive at a solution. A post-ICA stage is applied in which all ICs are adjusted to have positive peak values for enhanced visualisation, and the mixing matrix is also scaled so that the dominant relationships between the ICs and the spectral signatures can be easily identified.

The signs of the ICs are first manipulated so that the dominant spectrum-like peak in each component is positive, to make them resemble spectrum-like features. The algorithm to achieve the sign constraints is as follows:

1. find the sign of the maximum absolute value, i.e., the sign of the dominant peak, for every IC, y_j' , denoted by SN_j where $(j=1\dots n)$,
2. adjust the mixing matrix, $\underline{\mathbf{A}}$, and the IC matrix, \mathbf{Y} :

$$\mathbf{B} = \underline{\mathbf{A}} \text{diag}(SN_1, SN_2, \dots, SN_n)$$

$$\mathbf{S} = \text{diag}(SN_1, SN_2, \dots, SN_n)\mathbf{Y}$$

then

$$\mathbf{X} = \mathbf{BS}$$

where $\text{diag}(\cdot)$ is an $n \times n$ diagonal matrix.

A new term *significance index* has been created to underline the importance of the elements of the resulting mixing matrix. A significance index of 1 represents the strongest influence from an IC onto a power spectrum signature, whereas a smaller significance index implies that the influence is weaker. Note that there is no equivalent relationship in PCA.

The algorithm to achieve the unit upper bound of significance indices for each IC is equivalent to the adjustment of the maximum value of each column of the mixing matrix $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2 \cdots \mathbf{b}_n]$ to be unity:

1. find element of each column of \mathbf{B} with the maximum absolute value (i.e. the vector ∞ -norm of each column of \mathbf{B}):

$$\Delta_j = \|\mathbf{b}_j\|_{\infty}, j = 1 \cdots n$$

2. scale the mixing matrix \mathbf{B} and the IC matrix \mathbf{S} :

$$\mathbf{A} = \mathbf{B} \text{diag}(\Delta_1^{-1}, \Delta_2^{-1}, \cdots \Delta_n^{-1})$$

$$\mathbf{C} = \text{diag}(\Delta_1, \Delta_2, \cdots \Delta_n) \mathbf{S}$$

then

$$\mathbf{X} = \mathbf{AC} = [\mathbf{a}_1, \mathbf{a}_2 \cdots \mathbf{a}_n] \begin{bmatrix} \mathbf{c}'_1 \\ \mathbf{c}'_2 \\ \vdots \\ \mathbf{c}'_n \end{bmatrix} = \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix} \mathbf{c}'_1 + \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix} \mathbf{c}'_2 + \cdots + \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix} \mathbf{c}'_n \quad (6.11)$$

where $\mathbf{a}_j = [a_{1,j}, a_{2,j} \cdots a_{m,j}]'$, $j=1 \cdots n$, is the j^{th} column of the scaled mixing matrix \mathbf{A} and \mathbf{c}'_j ($j=1 \cdots n$) is the j^{th} spectrum-like IC. It can be seen that each power spectrum or P_i ($i=1 \cdots m$) in \mathbf{X} is represented by a linear combination of the ICs. The j^{th} column of the scaled \mathbf{A} matrix relates to the j^{th} spectrum-like independent component \mathbf{c}'_j , and value $a_{i,j}$, i.e. the significance index, represents the influence from the j^{th} independent component \mathbf{c}'_j onto the power spectrum of the i^{th} sensor observations, i.e. the i^{th} process variable.

6.2.3 Some Practicalities

ICA methods do not provide a means of determining the number of independent components that should be included in the model. An initial condition of the mixing matrix is also required for the maximisation process. Principal component analysis (PCA) can be used to provide both of these. A truncated PCA model can have the same structure as that of equation (6.11) plus an error matrix:

$$\mathbf{X} = \begin{bmatrix} t_{1,1} \\ t_{2,1} \\ \vdots \\ t_{m,1} \end{bmatrix} \mathbf{v}'_1 + \begin{bmatrix} t_{1,2} \\ t_{2,2} \\ \vdots \\ t_{m,2} \end{bmatrix} \mathbf{v}'_2 + \dots + \begin{bmatrix} t_{1,n} \\ t_{2,n} \\ \vdots \\ t_{m,n} \end{bmatrix} \mathbf{v}'_n + \mathbf{E} = \mathbf{TV}' + \mathbf{E}, \quad (6.12)$$

i.e., rows of \mathbf{X} can be reconstructed from $\mathbf{v}'_1 \dots \mathbf{v}'_n$ orthogonal basis functions, where \mathbf{T} is a $m \times n$ matrix with elements $t_{i,j}$ ($i=1 \dots m, j=1 \dots n$), matrix $\mathbf{V}' = [\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_n]$ represents the dominant row-major principal components (PCs) and \mathbf{E} an error matrix, which includes truncated principal components. The decision to truncate is made when the eigenvalue associated with the next principal component represents less than $\alpha\%$ of the sum of all the eigenvalues. Parameter $\alpha\%$ is typically in the range from 0% to 5%. The number of dominant ICs is set to the same number of dominant PCs. The initial value of the mixing matrix of the ICA model is set to the matrix \mathbf{T} rather than a random default setting, when FastICA algorithm is applied. This customised initial value is found to be better than the random one in terms of the repetition of the ICA result.

6.2.4 Dominance of ICs and Isolation of Sources

Dominance of ICs

Spectral ICA decomposition will extract most of the dominant peaks provided the number of independent components is large enough. The period of each oscillation can be estimated from the peak position of the IC. The total energy of all the power spectra in \mathbf{X} can be approximated by $\|\mathbf{AC}\|_{sum}$, which can also be interpreted as the total energy carried by all dominant oscillatory ICs, where $\|\cdot\|_{sum}$ is the sum-norm of a matrix (i.e. the sum of the absolute values of all its elements). The total energy related to the j^{th} IC is $\|\mathbf{a}_j \mathbf{c}_j'\|_{sum}$. Each dominant independent component may be associated with a ratio, termed the Dominance Ratio (DR), which reflects the influence or dominance of that IC. Dominance is specified as the ratio of the total energy carried by that IC over the total energy carried by all dominant oscillatory ICs:

$$DR(j) = \frac{\|\mathbf{a}_j \mathbf{c}_j'\|_{sum}}{\|\mathbf{AC}\|_{sum}} = 100 \left\{ \frac{\|\mathbf{a}_j \mathbf{c}_j'\|_{sum}}{\sum_{j=1}^n (\|\mathbf{a}_j \mathbf{c}_j'\|_{sum})} \right\} \% \quad (6.13)$$

The larger the $DR(j)$ value, the greater the influence of the j^{th} IC in the perturbed plant. An IC is said to be dominant if its DR value is not too small, where the threshold depends on the number of peaks of interest and is typically 0.2%. Thus the new method also gives the period and energy percentage, i.e., DR value, for each IC.

Isolation of Sources

Each dominant IC is assumed to represent a separate source. Its propagation around the plant can be found by looking at the significant indices, either in tabular form or as a plot. For the j^{th} dominant IC, a large significance index value of $a_{i,j}$, in the j^{th} column of the **A** matrix, represents a significant linkage between the j^{th} IC and the i^{th} power spectrum, and hence the i^{th} process variable. In particular, a value of $a_{i,j} = 1$ suggests that the source of the j^{th} IC oscillation, i.e. the root cause of this oscillation, is most probably closest to the i^{th} process variable. If this oscillation was to propagate to the k^{th} process variable, resulting in a secondary oscillation in that process variable, its significance index $a_{k,j}$ will also be significant, i.e. a lot bigger than 0. A significance index in the k^{th} process variable close to 1 would indicate a strong direct link between the i^{th} and k^{th} process variables. Examples of such deductions are given in section 6.3.2. Significance indices associated with one or more ICs are compared by plotting them as stacked bar charts like that shown in Figure 6-4. Each bar represents a separate process variable and is formed on the basis of the contributing ICs, where the total length is the sum of the contributing IC significance indices.

6.2.5 Multi-Range Spectral ICA

Experience has shown that some oscillations in either the middle frequency or high frequency ranges might not be detectable, because oscillations at low frequencies can dominate the total energy. In order to improve the detection resolution in the middle frequency to high frequency ranges, a multi-range spectra ICA is preferred. An automated way of choosing frequency channels for different ranges was

described in (Thornhill *et al.*2003) by frequency domain filtering. The multi-range spectral ICA is achieved by three steps:

Step 1: the power spectra are partitioned into 3 different frequency ranges: low frequency, middle frequency and high frequency range, i.e.,

$$\mathbf{X} = \left(\underbrace{\begin{pmatrix} P_1(f_1) & \cdots & P_1(f_{N_1}) \\ \cdots & \cdots & \cdots \\ P_m(f_1) & \cdots & P_m(f_{N_1}) \end{pmatrix}}_{\text{Low-f, } \mathbf{X}_1} \underbrace{\begin{pmatrix} P_1(f_{N_1+1}) & \cdots & P_1(f_{N_2}) \\ \cdots & \cdots & \cdots \\ P_m(f_{N_1+1}) & \cdots & P_m(f_{N_2}) \end{pmatrix}}_{\text{Mid-f, } \mathbf{X}_2} \underbrace{\begin{pmatrix} P_1(f_{N_2+1}) & \cdots & P_1(f_N) \\ \cdots & \cdots & \cdots \\ P_m(f_{N_2+1}) & \cdots & P_m(f_N) \end{pmatrix}}_{\text{High-f, } \mathbf{X}_3} \right)$$

where f_{N_1} and f_{N_2} separate the three frequency ranges: $0 < f_{N_1} < f_{N_2} < f_N$. The clarified description could be $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3]$. Before applying spectral ICA on each of the sub-set $\mathbf{X}_1, \mathbf{X}_2$ and \mathbf{X}_3 , each row of each sub-set matrix should be normalised as described in equation (6.8). The normalised version is then:

$$\bar{\mathbf{X}}_k = \mathbf{\Omega}_k^{-1} \mathbf{X}_k \quad (k=1,2,3) \quad (6.14)$$

where $\mathbf{\Omega}_k$ is $m \times m$ diagonal matrix with the element value at position (m,m) being equal to the sum of the m^{th} row of \mathbf{X}_k .

Step 2: basic spectral ICA is performed on each of the three normalised data sub-sets. Each analysis will result in n_k ICs:

$$\bar{\mathbf{X}}_k = \mathbf{A}_k \mathbf{C}_k \quad (k=1,2,3) \quad (6.15)$$

where \mathbf{A}_k is the $m \times n_k$ mixing matrix and \mathbf{C}_k is the $n_k \times r_k$ IC matrix, where $r_1 = N_1, r_2 = N_2 - N_1$ and $r_3 = N - N_2$. Substitute equation (6.15) into equation (6.14) to get:

$$\mathbf{X}_k = \mathbf{\Omega}_k \mathbf{A}_k \mathbf{C}_k \quad (k=1,2,3) \quad (6.16)$$

Step 3: the results are combined, ensuring that the final result has the same form as equation (6.11), by padding out matrices with zeros. The alignment is made on both \mathbf{X}_k and \mathbf{C}_k :

$$\underbrace{\begin{bmatrix} \mathbf{X}_1, \mathbf{0}_{m \times (N-N_1)} \end{bmatrix}}_{\mathbf{Z}_1} = \mathbf{\Omega}_1 \mathbf{A}_1 \underbrace{\begin{bmatrix} \mathbf{C}_1, \mathbf{0}_{n_1 \times (N-N_1)} \end{bmatrix}}_{\mathbf{E}_1} \quad (6.17)$$

$$\underbrace{\begin{bmatrix} \mathbf{0}_{m \times N_1}, \mathbf{X}_2, \mathbf{0}_{m \times (N-N_2)} \end{bmatrix}}_{\mathbf{Z}_2} = \mathbf{\Omega}_2 \mathbf{A}_2 \underbrace{\begin{bmatrix} \mathbf{0}_{n_2 \times N_1}, \mathbf{C}_2, \mathbf{0}_{n_2 \times (N-N_2)} \end{bmatrix}}_{\mathbf{E}_2} \quad (6.18)$$

$$\underbrace{\begin{bmatrix} \mathbf{0}_{m \times N_2}, \mathbf{X}_3 \end{bmatrix}}_{\mathbf{Z}_3} = \mathbf{\Omega}_3 \mathbf{A}_3 \underbrace{\begin{bmatrix} \mathbf{0}_{n_3 \times N_2}, \mathbf{C}_3 \end{bmatrix}}_{\mathbf{E}_3} \quad (6.19)$$

The synthetic result will then be:

$$\mathbf{X} = \mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3 = \underbrace{\begin{bmatrix} \mathbf{\Omega}_1 \mathbf{A}_1, \mathbf{\Omega}_2 \mathbf{A}_2, \mathbf{\Omega}_3 \mathbf{A}_3 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \\ \mathbf{E}_3 \end{bmatrix}}_{\mathbf{C}} = \mathbf{AC}, \quad (6.20)$$

which is exactly the same model as that of the basic spectral ICA. The post-ICA algorithm for determination of significance indexes should be performed on this synthetic result.

6.3 Examples

Section 6.3 contains two examples. The first uses simulated data to illustrate the basic features of spectral ICA and the second is a case study that applies multi-range spectral ICA to a real plant data obtained courtesy of the Eastman Chemical Company, Kingsport, Tennessee.

6.3.1 Example1: Simulated Plant Data

The simulated plant (Figure 6-1) contains two interacting, PI control loops, Loop1 and Loop2 that were optimally tuned and excited by coloured noise. Independent oscillatory disturbances are introduced into both loops during steady state. Their periods are 200 samples/cycle (Loop 1) and 360 samples/cycle (Loop 2). Both the controlled process variables (PVs) and the controller outputs (OPs) were recorded, resulting in four 4000 sample records. The normalised records and associated power spectra are shown in Figure 6-2 where the frequency axis in the right hand panel is normalised by the sampling frequency. Note that a value of \bar{f} on the normalised frequency axis represents $1/\bar{f}$ samples per cycle, e.g. $\bar{f} = 0.01$ is equivalent to 100 samples/cycle. According to the Shannon sampling theorem, the normalised frequency range for one-sided spectra cannot exceed the Nyquist frequency i.e. the range ≤ 0.5 or 2 samples/cycle.

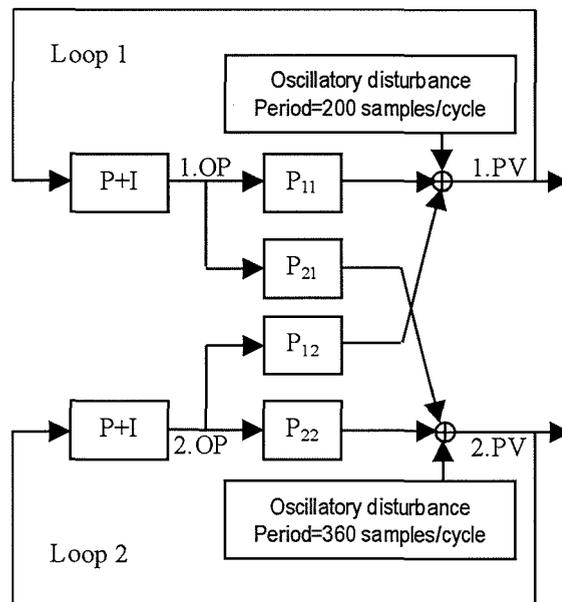


Figure 6-1: The simulation plant schematic

It can be seen from the right-hand panel of Figure 6-2 that each power spectrum contains two dominant peaks, which represent two dominant oscillatory disturbances. Spectral ICA is therefore required to extract two ICs, one IC should contain one of the two dominant peaks at 200 samples/cycle (0.005 on the normalised frequency axis), whilst the other IC should contain the other at 360 samples/cycle (0.00278 on the normalised frequency axis). The sources of these 2 disturbances should also be correctly isolated to the relevant loops by means of the significance indices. Both basic spectral ICA and spectral PCA were performed on these records. The two dominant ICs and related principal components (PCs) are shown in Figure 6-3 ($\alpha\% = 1\%$ was used in this case). Values of the significance indices (i.e. the a_{ij} values), the dominance ratio (DR) and estimated periods are listed in Table 6-1.

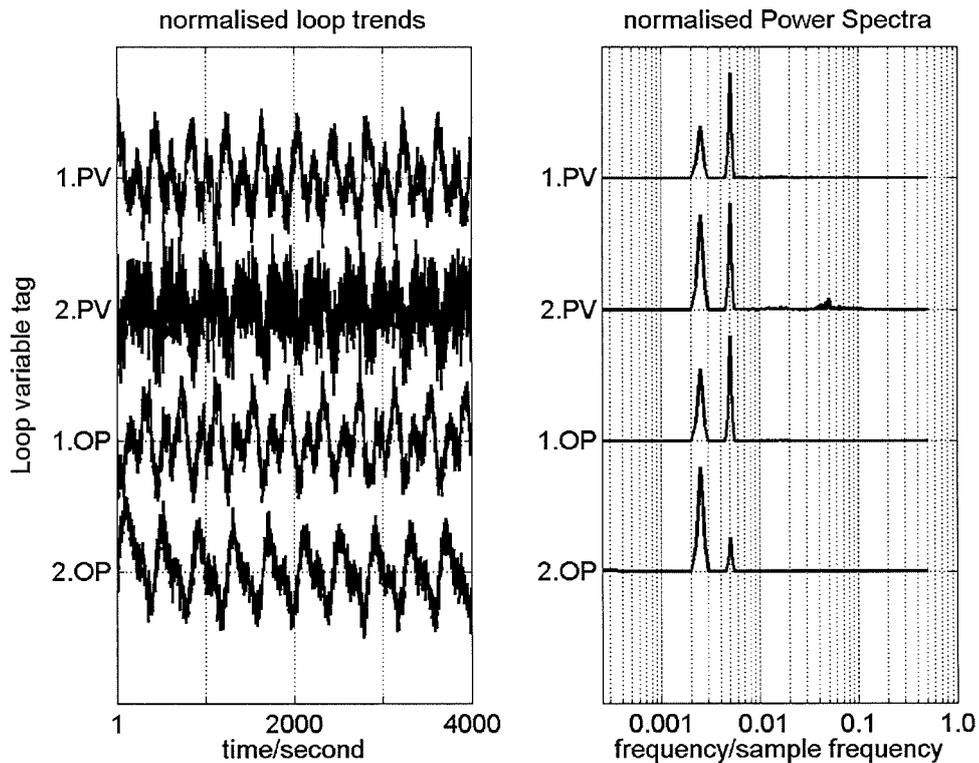


Figure 6-2: Normalised loop trends and normalised power spectra

Detection of two oscillatory disturbances

The dominance ratio values for the two ICs in Table 6-1 were 52.7% for IC1 and 47.3% for IC2, indicating that both IC1 and IC2 are dominant because each of them represents nearly half of the total energy. Each of the two spectrum-like dominant ICs in the left-hand panel of Figure 6-3 has a single sharp peak indicating an oscillatory data. The period results in Table 6-1 show that IC1 contains a peak located at 200 samples/cycle (0.005 on the normalised frequency axis) and IC2 contains a peak at 364 samples/cycle (0.00275 on the normalised frequency axis). These two periods match the periods of the real disturbances to the accuracy of the discretization process.

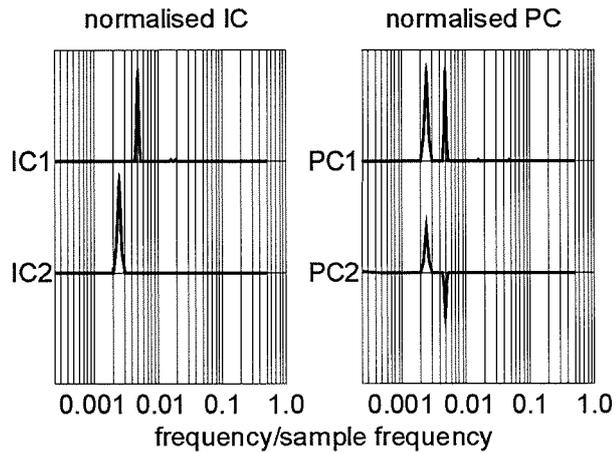


Figure 6-3: Independent components (ICs) and principal components (PCs) of example 1

significance index	IC1	IC2
	Loop1.PV	1.00
Loop2.PV	0.35	0.27
Loop1.OP	0.92	0.55
Loop2.OP	0.37	1.00
DR	52.7%	47.3%
Period(samples/cycle)	200	364

Table 6-1: Values of significance indexes, dominance ratios and estimated periods for example 1

Isolation of the two oscillatory disturbances

The significance indices are tabulated in Table 6-1. Loop1 is correctly isolated as the source of the oscillatory disturbance represented by IC1, because Loop1.PV has a significance index of 1.00 for IC1. Similarly, Loop2 is correctly isolated as the source of IC2, because Loop2.OP has a significance index of 1.00 for IC2. It is interesting to note that the index for Loop2.PV is relatively small, indicating that this loop is tightly regulated and the effect is therefore more visible in Loop2.OP. That is, the controller is compensating for the disturbance by means of movements

in the controller output. All the significance indices for both ICs are large because the loops are tightly coupled, as can be seen also in Figure 6-2, which shows that all the variables contained both spectral peaks in varying degrees.

The significance index plot for both ICs (Figure 6-4) gives a visual way to isolate the sources. The source of a particular IC can be associated with that loop variable tag that has a bar of unit length. For example, loop variable tag 1.PV in Figure 6-4 has a unit-length bar associated with IC1, whilst tag 2.OP has a unit-length bar associated with IC2. Note that the legend of a significance index plot indicates not only the discrimination of the ICs, but also their periods and total energy percentages from the dominance ratios.

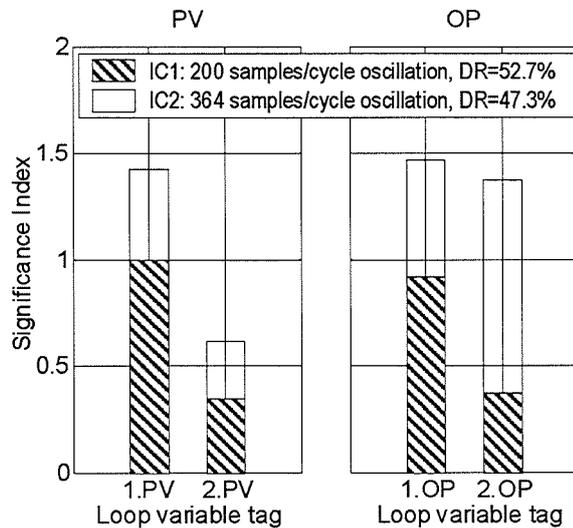


Figure 6-4: Significance index plot for IC1 and IC2 of example 1

Comparison of spectral ICA and spectral PCA

The two dominant principal components (PCs) are shown in the right-hand panel of Figure 6-3. Both PC1 and PC2 contain two peaks. They are orthogonal but contain more than one frequency component. Single peaks resembling the ICs in the left hand panel of Figure 6-3 can be created by linear combinations of PC1 and PC2, for instance a weighted sum would achieve cancellation of the right-hand peak while a weighted difference would achieve cancellation of the left hand peak. That observation illustrates why PCA forms a good starting point for ICA. PCA identifies the dominant spectral peaks and the ICA then separates them. The interpretation of detection, isolation and grouping of oscillation disturbances based on spectral PCs is not straightforward because the same spectral peaks are present in more than one PC and becomes difficult when the number of dominant PCs increases. On the other hand, ICA focuses on the extraction of spectrum-like components with single peaks, such as IC1 and IC2 in Figure 6-3. The power spectrum of a loop trend can then be approximated by a simple combination of the two ICs. The significance index comparison for each IC gives a natural and straightforward way to isolate real disturbances.

6.3.2 Example2: Case Study on the Industrial Data Set

The industrial data set was provided courtesy of the Eastman Chemical Company, Kingsport, Tennessee. It consists of 30 process variables (PVs) including 15 controlled variables and 15 uncontrolled variables from indicators. Controller output (OP) records were available for each controller. Thus an OP tag number will be the same as its associated controlled PV tag number. These tag numbers are also used to identify the 15 control loops. The time trends and their power spectra are shown in Figure 6-5 and Figure 6-6 respectively. Gaps in the tag numbering for the OPs indicate that some tags are not associated with controllers. The sampling interval was 20s and data records of 8192 samples were used to generate the power spectra by FFT with $\tilde{N} = 8192$. Thus the rows of the data matrix \mathbf{X} had 4096 elements up to the Nyquist frequency.

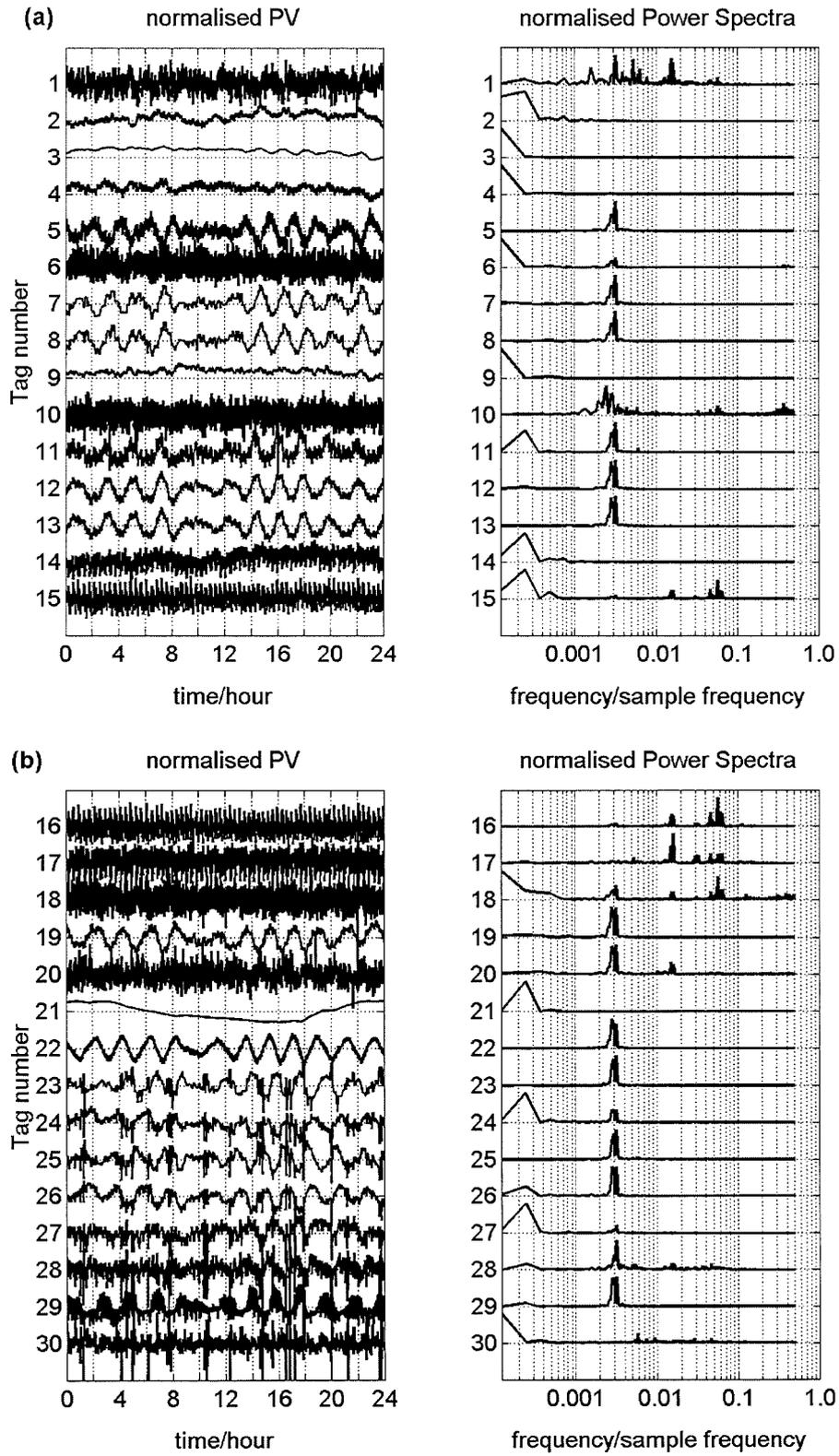


Figure 6-5: Normalised PVs and power spectra

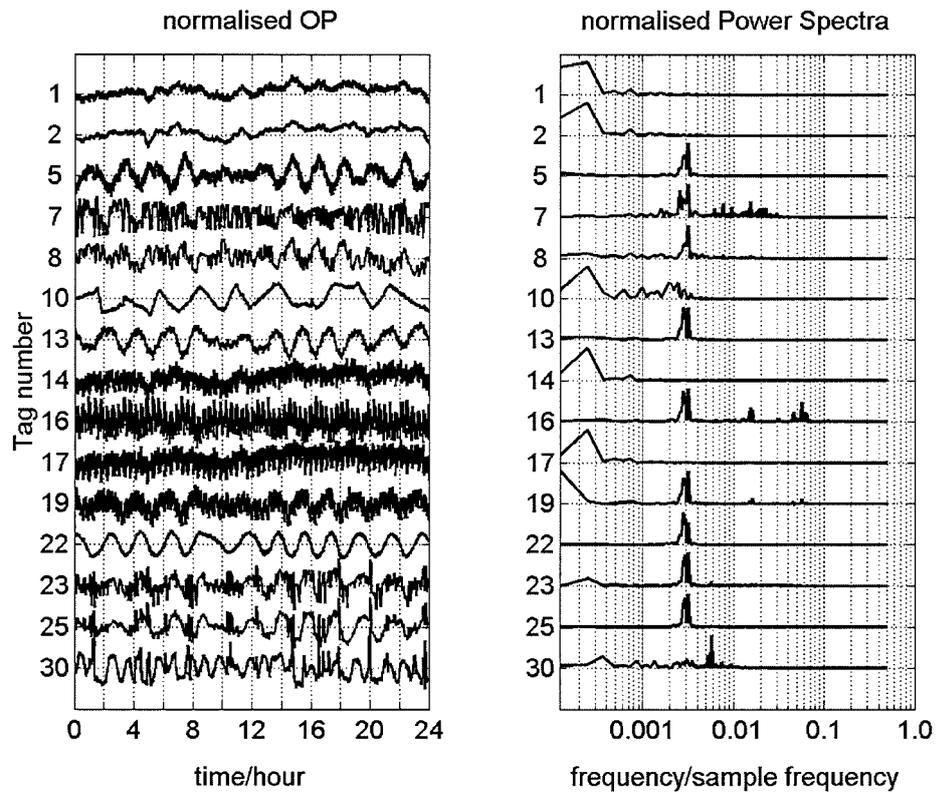


Figure 6-6: Normalised OPs and power spectra

Detection of oscillatory disturbances

Figure 6-5 and Figure 6-6 show that many of the PV and OP spectra contain sharp peaks that correspond to oscillatory trends in the time domain. It is clear that there is a dominant, plant-wide oscillation corresponding to peaks at about 0.0028 to 0.0032 on the normalised frequency axis (356 to 315 samples/cycle). This oscillation had already been examined and a sticking valve was found in the loop with Tag 22 (Thornhill *et al.* 2002b). Spectral ICA should detect and isolate not only this dominant oscillatory source, but also other oscillations in the entire frequency range.

Multi-range spectral ICA was applied because some of the less dominant peaks such as those in Tags 16 to 18 appear in the medium to high frequencies. Table 6-2 lists the frequencies that were chosen to separate the three frequency ranges and the parameters $\alpha\%$, which can be different for each range. The smaller $\alpha\%$ value in the low frequency range was used to ensure a higher resolution, because more narrow-band peaks were found in the low frequency range. Figure 6-7 shows the results of ten ICs (IC1 to IC5 in the low frequency range, IC6 to IC8 in the middle frequency range and IC9 to IC10 in the high frequency range). The dominance ratio and estimated periods are listed in Table 6-3. Significance index plots for different frequency ranges are shown in Figure 6-8 and Figure 6-9.

f_{N_1}		0.01 (100 samples/cycle)
f_{N_2}		0.025 (40 samples/cycle)
$\alpha\%$	Low- f range	0.4%
	Mid- f range	4%
	High- f range	4%

Table 6-2: Separation of the multi-frequency ranges and their associated $\alpha\%$ parameters

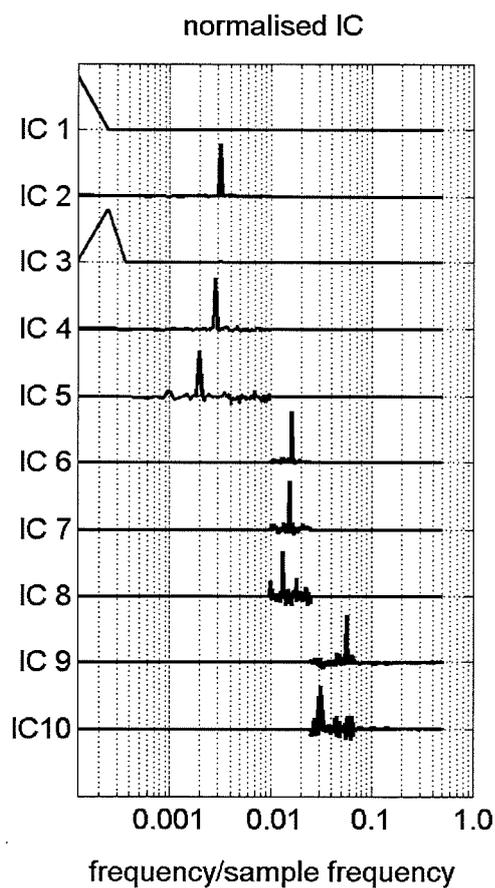


Figure 6-7: Independent components of example 2

	DR	Period (samples/cycle)
Low- <i>f</i> range		
IC1	14.0%	8192
IC2	23.5%	315
IC3	14.7%	4096
IC4	23.9%	356
IC5	2.1%	512
Mid- <i>f</i> range		
IC6	0.4%	62
IC7	0.8%	66
IC8	0.2%	76
High- <i>f</i> range		
IC9	16.2%	18
IC10	4.2%	32

Table 6-3: Dominance ratio (DR) and period results of example 2

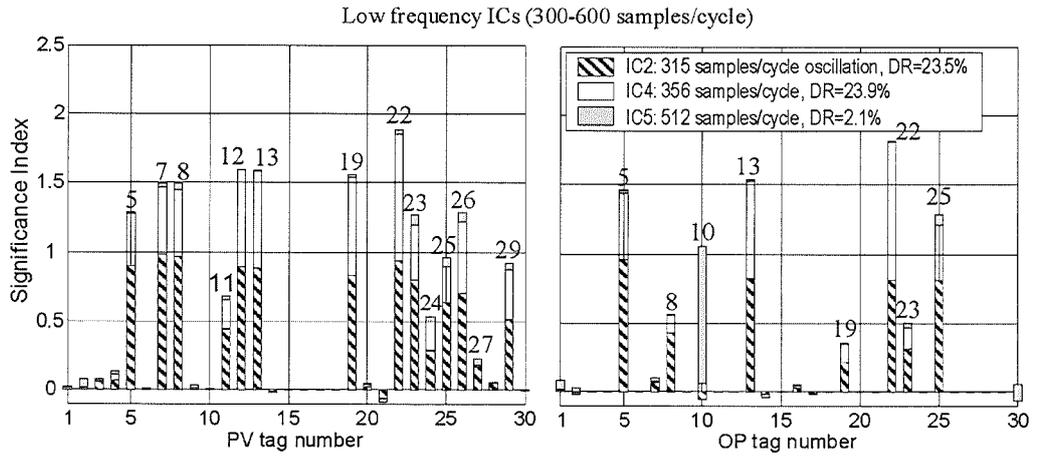


Figure 6-8: Significance index plot for low frequency IC2, IC4 and IC5

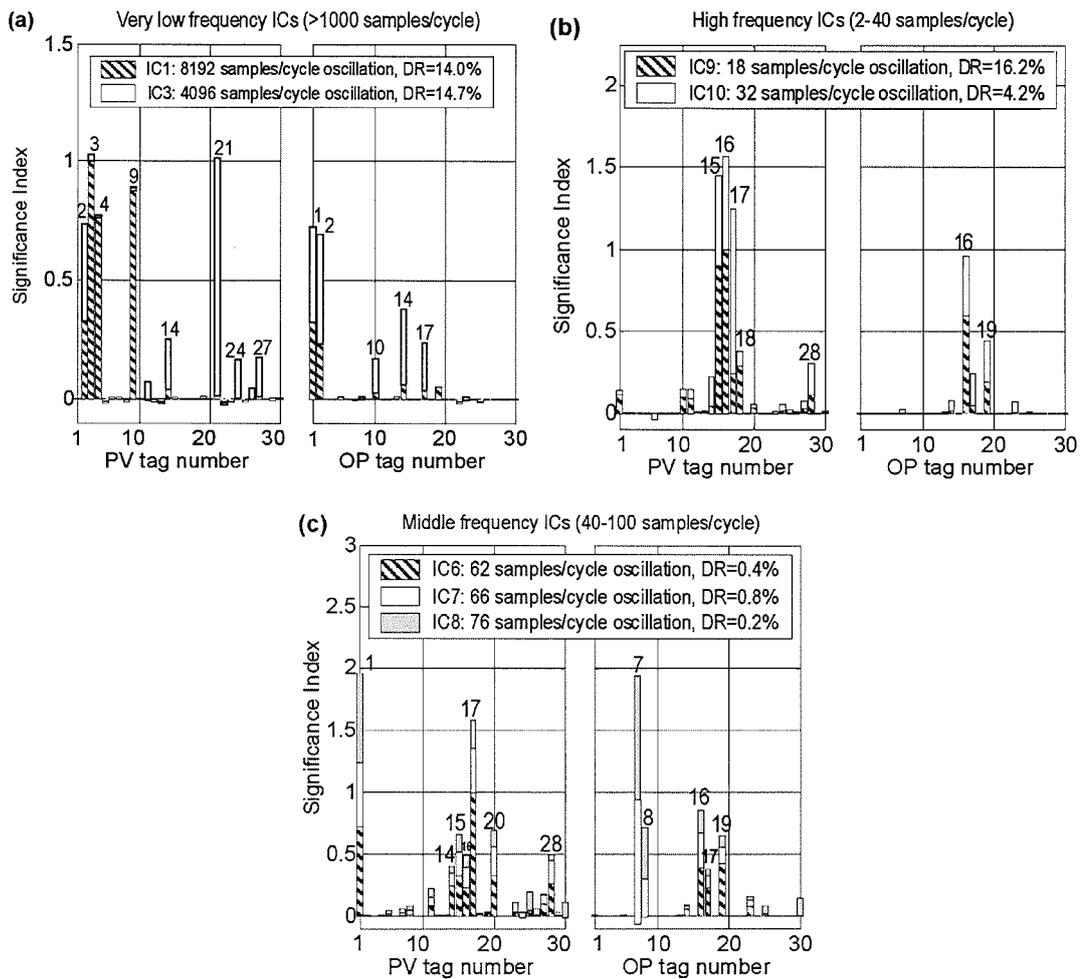


Figure 6-9: Other significance index plots for ICs in different frequency ranges

Each detected IC in Figure 6-7 contains a single dominant peak, which indicates an oscillatory source in the data. It can be seen from Table 6-3 that:

- IC2 and IC4 should be examined first because they represent 47.4% of the total energy (with dominance ratios of 23.5% + 23.9%); they contain oscillatory sources with periods of 315 and 356 samples/cycle;
- IC1 and IC3 are very low frequency components that represent 28.7% of the total energy percentages (14.0% + 14.7%); they contain oscillatory sources with periods of 8192 and 4096 samples/cycle;
- IC5 with an energy of 2.1% is small but not negligible. It also should be investigated. The period of oscillation is 512 samples per cycle;
- IC9 and IC10 are high frequency components that represent 20.4% of the total energy (16.2% + 4.2%); they contain oscillatory sources with periods of 18 and 32 samples/cycle; IC9 is most dominant in this frequency range;
- IC6, IC7 and IC8 are middle frequency components with periods of 62, 66 and 76 samples/cycle, and contain only 1.4% of the total energy (0.4% + 0.8% + 0.2%). In spite of their lack of dominance, they are still detectable by means of multi-range analysis.

Isolation of the oscillatory sources

Isolation of these oscillatory sources is based on an examination of the significance index plots. It is assumed that a large significance index for a given tag means that the IC in question is closely associated with that tag, which is therefore a candidate for the source.

Figure 6-8 shows the low frequency components IC2, IC4 and IC5. The figure suggests that Tag 22 is the source of IC2 and IC4 (315 and 356 samples/cycle) and that those oscillations are linked together and have propagated widely. In fact, tag 22 was known to have a faulty control valve that was causing a limit cycle oscillation. The reasoning behind these conclusions is as follows:

- the most dominant oscillatory components, i.e. IC2 and IC4, always appear together in almost all of the perturbed process measurements (both PVs and OPs);
- IC2 and IC4 appear in 21 tags;
- tags 22.PV and 22.OP have the biggest significance indices for the combination of IC2 and IC4.

The conclusion that the IC2 and IC4 oscillations are linked may also be inferred from inspection of the original spectra. Many of the spectra have a double peak at about 0.003 on the frequency axis. The 315 and 356 samples/cycle oscillations are the two components within that double peak. These two oscillations can also be seen in the time domain plots. For instance, in Tag 22 the peaks are closer together in the time period from 14-18 hours and wider apart between 18 to 24 hours. It is possible that the plant operators working on different shifts chose to run the plant with slightly different settings.

Further observations focus on the low energy oscillation in IC5 at 512 samples per cycle:

- Tag 10.OP is different from others, it has a big significance index value for IC5;
- no other process measurements are significant for IC5.

These findings show that the oscillation of IC5 (512 samples/cycle) is located only in the control loop of Tag 10, and this oscillation doesn't propagate. Moreover, it is present only in the OP of Tag 10 and not in the PV thus suggesting the control loop for Tag 10 is successfully rejecting an external disturbance. This demonstrates the importance of analysing OP data.

Figure 6-9(a) shows the very low frequency components IC1 and IC3. Both can be assigned to a source and both have propagated:

- the source of IC1 most directly links to the process variable 3.PV, because of its largest significance index for IC1. Process measurements with PV tags 2, 4 and 9 and OP tags 1 and 2 are also influenced by this slow oscillatory source;
- the source of IC3 most directly links to the process variable 21.PV, and its effect can also be seen in the data pertaining to PV tags 2, 14, 24 and 27 and OP tags 1, 2, 10, 14 and 17.

Figure 6-9(b) shows the high frequency components IC9, with 18 samples per cycle, and IC10, with 32 samples per cycle. IC9 is a dominant oscillation having a dominance ratio of 16.2%. Its significance index suggests it is associated with Tag

16. The oscillation with 32 samples per cycle (IC10) is most strongly linked to Tag 17.
17. Both oscillations have propagated to a few other locations (e.g. tags 15 and 18).

Figure 6-9(c) shows the middle frequency components IC6, IC7 and IC8. These middle frequency oscillations are less dominant because they have small dominance ratio (DR) values. Isolation of these sources is still feasible by means of the significance indices, as follows:

- the control loop for Tag 17 is responsible for the IC6 oscillation (see 17.PV);
- the control loop for Tag 7 is responsible for the IC7 and IC8 oscillation (see 7.OP).

By combining these observations it can be inferred that Tag 17 is the possible source of two separate oscillations, IC10 at 32 samples per cycle and IC6 at 62 samples per cycle. Given that 32 and 62 are close to having a ratio of 2:1 it may be the case that the IC10 oscillation is a harmonic of a non-sinusoidal oscillation at about 62 samples per cycle.

Some of the significance indices in Figure 6-8 and Figure 6-9 are very slightly negative, so slightly negative that they are unlikely to influence any conclusions that would be drawn. They arise because ICA performs a search resulting in independent components, which can have both negative as well as positive parts (e.g. see Figure 6-7) and hence significance indices can be both positive and negative. The sign correction algorithm described in Section 6.2.2 does not correct for the negative noise effects because it focuses on the dominant peaks.

6.4 Conclusions

The proposed spectral independent component analysis (spectral ICA) is based on the ICA analysis of spectra derived via a discrete Fourier transform from time domain process data. Spectral ICA is able to extract dominant spectrum-like independent components each of which has a narrow-bank peak that captures the behaviour of one of the oscillation sources. Proofs have been presented to underpin the theory of spectral ICA, in particular it has been shown that, for a data set having multiple oscillatory sources, the extraction of independent components with narrow-bank spectral peaks can be guaranteed by an ICA algorithm that maximises the kurtosis of the independent components (ICs).

The periods of the detected oscillations may be determined from the positions of the spectral peaks in the ICs and two indexes have been defined and shown to be effective in the isolation of the oscillatory sources. One is termed the dominance ratio (DR): it indicates the severity of an oscillation as a percentage of the total power. The other is a significance index that isolates an oscillatory source to a particular location in the plant. A bar chart is recommended to aid visualisation and comparison of the significance indexes. The benefits of these two indexes are that they help plant control engineers and maintenance personnel to prioritise the faults and to focus effort in the right part of the plant.

This novel method has been applied to a simulated data set and to plant data obtained from an industrial chemical plant. Results demonstrate its ability to detect and isolate multiple dominant oscillations in different frequency ranges. The introduction of controller output actuator signals (OPs) as well as plant

measurements (PVs) into the analysis was found to be important for the achievement of higher accuracy of source isolation. For instance, it was possible to demonstrate that one oscillatory source was being successfully compensated for by a controller because the significance index was much larger for the OP than for the PV.

The starting point for spectral ICA is a spectral principal component analysis (spectral PCA). The benefit to be gained from the additional computational steps of spectral ICA is that the ICs each contain one narrow band peak compared to spectral PCA where the principal components may contain more than one peak. The isolation of the oscillatory sources is thus aided and enhanced by adding a spectral ICA stage to a multivariate spectral analysis.

Chapter 7

Conclusions and Recommendations

7.1 Conclusions

It is important for process control engineers to detect and diagnose control loop problems in chemical process plants quickly as these can affect plant performance. A large petrochemical plant may have a 1000 or more control loops and indicators, so a key requirement of an industrial control engineer is for an automated means to detect and isolate the root causes so that maintenance effort can be directed efficiently. Detection and diagnosis of plant-wide oscillations are of particular importance, because the propagation of an oscillation throughout the plant can have an impact on product quality and running costs.

Although there has been considerable commercial and academic interest in methods, as described in Chapter 2, for analysing the performance of control systems based on the minimum variance control, they are usually designed to detect, and not to isolate or diagnose, faults or other root causes of poor performance. Chapter 2 also reviewed some attempts that have been made to enhance performance assessment with methods to detect and diagnosis oscillations in control loops, however the isolation and diagnosis of multiple oscillations has not been fully studied in a multi-loop arrangement. This work addresses these deficiencies, and seeks to solve these problems.

Chapter 3 proposed various statistics to facilitate the status monitoring of PI/PID loops. A quantitative version of loop status (*LS*) can be formulated by assigning different real number values to the various categories and a filtered & smoothed statistic (*LL*) can be obtained by applying an EWMA filter. The loop status indicates what kind of deterministic trend is present in a PI/PID loop and, if abnormal, what is the possible cause. Although the approach has been developed for PI/PID controllers, it should be equally applicable for those controllers that have a frequency dependent statistic *R*, which is the ratio of the signal to noise ratios of the controlled variable and controller output. Thus the approach is not suitable for a P control loop, because *R* will always be unity. Equally it is also not feasible for open-loop control, because *R* is intended to reflect the relationship between the controlled variable and the controller output, and this does not exist for an open-loop control system. Loop status could have a number of different uses:

- as an operator aid; a loop status icon on a plant schematic might provide an overall impression of control activity on a plant;
- as an aid for the control/maintenance engineer: Table 3-1 lists some common causes of Loop Statuses;
- as a pre-processing procedure for the localisation/isolation of the source loop of a fault/disturbance in a unit-wide or plant-wide control system containing a number of interacting loops (Chapter 4).

Loop status can form the basis for the development of a fault localisation technique based on the Overall Loop Performance Index (OLPI) (Chapter 4). Fault localisation based on OLPI comparison has proved to be both successful and robust in a number of typical plant simulations and when analysing sets of real industrial

data. OLPI comparison can point out a problem loop in a multi-loop arrangement, and Loop Status information of that problem loop helps to narrow down possible root causes. This approach is successful when analysing a plant with interacting loops that is perturbed by ONE dominant deterministic disturbance at a time.

However the OLPI-based technique is unlikely to work for either multiple faults that result in multiple dominant disturbances, or for poorly designed tightly-coupled plants.

A problem that is not often talked about is that of poor performance caused by tight-coupled loops due to poor plant design (Chapter 5). A data-driven diagnosis method for this problem is proposed. Although fairly simple, case study tests show its capability to distinguish between a poor design problem and a commissioning-stage fault, and to reason about poor process design.

An important aspect is the detection and isolation of multiple oscillations (Chapter 6). Spectral independent component analysis (spectral ICA) is based on the ICA of spectra obtained by discrete Fourier transforming process data. Spectral ICA is able to extract dominant spectrum-like independent components that capture the behaviour of oscillatory sources. Proofs have been presented to underpin the theory of spectral ICA, and two indexes have been defined and shown to be effective in the isolation of the oscillatory sources. The benefits of these two indexes are that they help plant control engineers and maintenance personnel to prioritise the faults and to focus effort in the right part of the plant. This novel method has been applied to a simulated data set and to plant data obtained from an industrial chemical plant.

Results demonstrate its ability to detect and isolate multiple dominant oscillations in different frequency ranges.

7.2 Future Work

Further research is needed to make ICA-based techniques more generally applicable:

- preliminary research has shown that the technique is able to isolate deterministic disturbances, such as drifts and non-stationary disturbances. This is likely to be more robust than loop status monitoring.
- Some advanced numerical issues need to be addressed in order to enhance the interpretation, for instance to identify those loops that contribute most strongly to a root cause. Non-negative matrix decomposition and non-negative sparse coding techniques are specific examples.
- The auto-covariance function of a data set has the same oscillatory behaviour as the raw data but the averaging operations in its calculation reduce the impact of noise. Performing ICA on the auto-covariances of plant measurements is likely to improve noise suppression and that avenue will be vigorously explored.
- Greater resolution and automation can be obtained by combining wavelet techniques with ICA.

Reference

- Akbaryan, F. and Bishnoi, P. R., (2001). Fault diagnosis of multivariate systems using pattern recognition and multisensor data analysis technique, *Computers & Chemical Engineering*, vol. 25, no. 9-10, pp. 1313-1339.
- Alexander, S. M. and Gor, T. B., (1998). Monitoring, diagnosis and control of industrial processes, *Computers & Industrial Engineering*, vol. 35, no. 1-2, pp. 193-196.
- Amand, T., Heyen, G., and Kalitventzeff, B., (2001). Plant monitoring and fault detection: Synergy between data reconciliation and principal component analysis, *Computers & Chemical Engineering*, vol. 25, no. 4-6, pp. 501-507.
- Aubrun, C., Robert, M., and Cecchin, T., (1995). Fault detection in a control loop, *Control Engineering Practice*, vol. 3, no. 10, pp. 1441-1446.
- Bakshi, B. R. and Stephanopoulos, G., (1994). Representation of process trends--IV. Induction of real-time patterns from operating data for diagnosis and supervisory control, *Computers & Chemical Engineering*, vol. 18, no. 4, pp. 303-332.
- Balle, P., (1999). Fuzzy-model-based parity equations for fault isolation, *Control Engineering Practice*, vol. 7, no. 2, pp. 261-270.
- Balle, P. and Fuessel, D., (2000). Closed-loop fault diagnosis based on a nonlinear process model and automatic fuzzy rule generation, *Engineering Applications of Artificial Intelligence*, vol. 13, no. 6, pp. 695-704.
- Basseville, M., (1997). Information Criteria for Residual Generation and Fault Detection and Isolation, *Automatica*, vol. 33, no. 5, pp. 783-803.
- Becraft, W. R. and Lee, P. L., (1993). An integrated neural network/expert system approach for fault diagnosis, *Computers & Chemical Engineering*, vol. 17, no. 10, pp. 1001-1014.
- Bendat, J. S. and Piersol, A. G., (1993). *Engineering applications of correlation and spectral analysis*, Second Edition, John Wiley & Sons.

- Berton, A. and Hodouin, D., (2003). Linear and bilinear fault detection and diagnosis based on mass and energy balance equations, *Control Engineering Practice*, vol. 11, no. 1, pp. 103-113.
- Bezergianni, S. and Georgakis, C., (2000). Controller performance assessment based on minimum and open-loop output variance, *Control Engineering Practice*, vol. 8, no. 7, pp. 791-797.
- Bissessur, Y., Martin, E. B., and Morris, A. J., (1999). Monitoring the performance of the paper making process, *Control Engineering Practice*, vol. 7, no. 11, pp. 1357-1368.
- Bloch, G., Ouladsine, M., and Thomas, P., (1995). On-line fault diagnosis of dynamic systems via robust parameter estimation, *Control Engineering Practice*, vol. 3, no. 12, pp. 1709-1717.
- Bogaerts, P., Cuvelier, A., and Kinnaert, M., (1998). Identification for fault detection in an industrial condenser, *Control Engineering Practice*, vol. 6, no. 10, pp. 1249-1256.
- Brydon, D. A., Cilliers, J. J., and Willis, M. J., (1997). Classifying pilot-plant distillation column faults using neural networks, *Control Engineering Practice*, vol. 5, no. 10, pp. 1373-1384.
- Chan, C. W., Jin, H., Cheung, K. C., and Zhang, H. Y., (2001). Fault detection of systems with redundant sensors using constrained Kohonen networks, *Automatica*, vol. 37, no. 10, pp. 1671-1676.
- Chen, B. H., Wang, X. Z., and McGreavy, C., (1998). On-line Operational Support System for Faults Diagnosis in Process Plants, *Computers & Chemical Engineering*, vol. 22, no. 1, p. S973-S976.
- Chen, B. H., Wang, X. Z., Yang, S. H., and McGreavy, C., (1999). Application of wavelets and neural networks to diagnostic system development, 1, feature extraction, *Computers & Chemical Engineering*, vol. 23, no. 7, pp. 899-906.
- Chen, J. and Patton, R. J., (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Kluwer Academic Publishers.
- Chen, J., (2001). *Control system Based Loop and Process Monitoring*, Ph.D. Thesis, University of Glasgow.

- Chen, J. and Howell, J., (2001). A self-validating control system based approach to plant fault detection and diagnosis, *Computers & Chemical Engineering*, vol. 25, no. 2-3, pp. 337-358.
- Chen, J. and Howell, J., (2002). Towards distributed diagnosis of the Tennessee Eastman process benchmark, *Control Engineering Practice*, vol. 10, no. 9, pp. 971-987.
- Choudhury, M. A. A. S., Shah, S. L., and Thornhill, N. F., (2002). Detection and diagnosis of system nonlinearities using higher order statistics, 15TH IFAC World Congress, Barcelona.
- Commault, C., (1999). On the disturbed fault detection and isolation problem, *Systems & Control Letters*, vol. 38, no. 1, pp. 73-78.
- Comon, P., (1994). Independent component analysis-a new concept?, *Signal Processing*, vol. 36, pp. 287-314.
- Daiguji, M., Kudo, O., and Wada, T., (1997). Application of Wavelet Analysis to Fault Detection in Oil Refinery, *Computers & Chemical Engineering*, vol. 21, no. 1, p. S1117-S1122.
- Dash, S. and Venkatasubramanian, V., (2000). Challenges in the industrial applications of fault diagnostic systems, *Computers & Chemical Engineering*, vol. 24, no. 2-7, pp. 785-791.
- Delfosse, N. and Loubaton, P., (1995). Adaptive Blind Separation of Independent Sources: a Deflation Approach, *Signal Processing*, vol. 45, pp. 59-83.
- Desborough, L. and Harris, T. J., (1992). Performance assessment measures for univariate feedback control, *Can.J.Chem.Eng.*, vol. 70, pp. 1186-1197.
- Desborough, L. and Harris, T. J., (1993). Performance assessment measures for univariate feedforward/feedback control, *Can.J.Chem.Eng.*, vol. 71, pp. 605-616.
- Desborough, L., Nordh, P., and Miller, R., (2001). Process out of control, *Industrial computing*, August, pp. 52-55.
- Dingli, Y., Shields, D. N., and Disdell, K., (1996). A simulation study on fault diagnosis of a high-temperature furnace using a bilinear observer method, *Control Engineering Practice*, vol. 4, no. 12, pp. 1681-1691.

- Dobson, J. P. and Thornhill, N. F., (2002). Analysis of plant-wide disturbances using conditional entropy, IEEE International Conference on Control Applications, Glasgow.
- Dong, S. N., Chang, W. J., Yong, J. C., and En, S. Y., (1996). Operation-aided system for fault diagnosis of continuous and semi-continuous processes, *Computers & Chemical Engineering*, vol. 20, no. 6-7, pp. 793-803.
- Downs, J. J. and Vogel, E. F., (1993). A Plant-Wide Industrial Process Control Problem, *Computers and Chemical Engineering*, vol. 17, no. 3, pp. 245-255.
- Dunia, R. and Qin, S. J., (1998). Joint diagnosis of process and sensor faults using principal component analysis, *Control Engineering Practice*, vol. 6, no. 4, pp. 457-469.
- Edelmayer, A., Bokor, J., Szigeti, F., and Keviczky, L., (1997). Robust Detection Filter Design in the Presence of Time-varying System Perturbations, *Automatica*, vol. 33, no. 3, pp. 471-475.
- Edwards, C., Spurgeon, S. K., and Patton, R. J., (2000). Sliding mode observers for fault detection and isolation, *Automatica*, vol. 36, no. 4, pp. 541-553.
- Ender, D. B., (1993). Process control performance: not as good as you think, *Control Engineering*, vol. 40, no. 10, pp. 180-190.
- Fasolo, P. S. and Seborg, D. E., (1995). Monitoring and fault detection for an HVAC control system, *HVAC & R Research*, vol. 1, pp. 177-193.
- Fisz, M., (1963). *Probability Theory and Mathematical Statistics* John Wiley & Sons, Inc..
- Forsman, K. and Stratin, A., (1999). A new criterion for detecting oscillations in control loops, European Control Conference, CP8-3, Karlsruhe, Germany.
- Fourie, S. H. and de Vaal, P., (2000). Advanced process monitoring using an on-line non-linear multiscale principal component analysis methodology, *Computers & Chemical Engineering*, vol. 24, no. 2-7, pp. 755-760.
- Frank, P. M., (1990). Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy – A survey and some new results, *Automatica*, vol. 26, pp. 459-474.

- Frank, P. M. and Koppen-Seliger, B., (1997). New Developments Using AI in Fault Diagnosis, *Engineering Applications of Artificial Intelligence*, vol. 10, no. 1, pp. 3-14.
- Fuente, M. J. and Vega, P., (1999). Neural networks applied to fault detection of a biotechnological process, *Engineering Applications of Artificial Intelligence*, vol. 12, no. 5, pp. 569-584.
- Garcia, E. A., Koppen-Seliger, B., and Frank, P. M., (1995). A frequency domain approach to residual generation for the industrial actuator benchmark, *Control Engineering Practice*, vol. 3, no. 12, pp. 1747-1750.
- Garcia, E. A. and Frank, P. M., (1997). Deterministic nonlinear observer-based approaches to fault diagnosis: a survey, *Control Engineering Practice*, vol. 5, no. 5, pp. 663-670.
- Garcia, F. J., Izquierdo, V., de Miguel, L. J., and Peran, J. R., (2000). Fault-diagnostic system using analytical fuzzy redundancy, *Engineering Applications of Artificial Intelligence*, vol. 13, no. 4, pp. 441-450.
- Gertler, J., (1998). *Fault detection and diagnosis in engineering systems*, Marcel Dekker, New York.
- Goh, K. B., Spurgeon, S. K., and Jones, N. B., (2002). Fault diagnostics using sliding mode techniques, *Control Engineering Practice*, vol. 10, no. 2, pp. 207-217.
- Grimble, M. J., (2002). Controller performance benchmarking and tuning using generalised minimum variance control, *Automatica*, vol. 38, no. 12, pp. 2111-2119.
- Hamelin, F. and Sauter, D., (2000). Robust fault detection in uncertain dynamic systems, *Automatica*, vol. 36, no. 11, pp. 1747-1754.
- Harris, T. J., (1989). Assessment of closed loop performance, *Can.J.Chem.Eng.*, vol. 67, no. 10, pp. 856-861.
- Harris, T. J., Boudreau, F., and MacGregor, J. F., (1996). Performance Assessment of Multivariable Feedback Controllers, *Automatica*, vol. 32, no. 11, pp. 1505-1518.
- Harris, T. J., Seppala, C. T., and Desborough, L. D., (1999). A review of performance monitoring and assessment techniques for univariate and multivariate control systems, *Journal of Process Control*, vol. 9, no. 1, pp. 1-17.

- Harris, T. J. and Seppala, C. T., (2001). Recent developments in controller performance monitoring and assessment techniques, *Chemical Process Control VI*, Tucson, AZ.
- Hägglund, T., (1995). A control-loop performance monitor, *Control Engineering Practice*, vol. 3, no. 11, pp. 1543-1551.
- Hägglund, T., (1999). Automatic detection of sluggish control loops, *Control Engineering Practice*, vol. 7, no. 12, pp. 1505-1511.
- Henry, D., Zolghadri, A., Monsion, M., and Ygorra, S., (2002). Off-line robust fault diagnosis using the generalized structured singular value, *Automatica*, vol. 38, no. 8, pp. 1347-1358.
- Himmelblau, D. M., (1978). *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*, Elsevier Scientific Publishing Company, Amsterdam.
- Hofling, T., Pfeufer, T., Deibert, R., and Isermann, R., (1995). An observer and signal-processing approach to FDI for the industrial actuator benchmark test, *Control Engineering Practice*, vol. 3, no. 12, pp. 1741-1746.
- Hofling, T. and Isermann, R., (1996). Fault detection based on adaptive parity equations and single-parameter tracking, *Control Engineering Practice*, vol. 4, no. 10, pp. 1361-1369.
- Horch, A., (1999). A simple method for detection of stiction in control valves, *Control Engineering Practice*, vol. 7, no. 10, pp. 1221-1231.
- Horch, A. and Isaksson, A. J., (1999). A modified index for control performance assessment, *Journal of Process Control*, vol. 9, no. 6, pp. 475-483.
- Howell, J., Chen, J., and Zhang, J., The Tennessee Eastman Problem as a process monitoring benchmark, IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes – Safeprocess'97 Preprints, Hull, UK, pp. 222-227.
- Huang, B., Shah, S. L., and Kwok, E. K., (1997). Good, Bad or Optimal? Performance Assessment of Multivariable Processes, *Automatica*, vol. 33, no. 6, pp. 1175-1183.
- Huang, B. and Shah, S. L., (1999). *Performance assessment of control loops*, Springer.
- Huang, B., (2002). Minimum variance control and performance assessment of time-variant processes, *Journal of Process Control*, vol. 12, no. 6, pp. 707-719.

- Huang, Y., Gertler, J., and McAvoy, T. J., (2000). Sensor and actuator fault isolation by structured partial PCA with nonlinear extensions, *Journal of Process Control*, vol. 10, no. 5, pp. 459-469.
- Hurri, J., Gävert, H, Särelä, J., and Hyvärinen, A., (1998). The FastICA package for MATLAB. Retrieved from <http://www.cis.hut.fi/projects/ica/fastica/>.
- Hyvärinen, A. and Oja, E., (2000). Independent components analysis: algorithms and applications, *Neural Networks*, vol. 13, pp. 411-430.
- Hyvärinen, A., Karhunen, J., and Oja, E., (2001). *Independent components analysis*, John Willey & Sons, Inc..
- Isermann, R., (1984). Process fault detection based on modeling and estimation methods – A survey, *Automatica*, vol. 20, pp. 387-404.
- Isermann, R., (1997). Supervision, fault-detection and fault-diagnosis methods - an introduction, *Control Engineering Practice*, vol. 5, no. 5, pp. 639-652.
- Isermann, R. and Balle, P., (1997). Trends in the application of model-based fault detection and diagnosis of technical processes, *Control Engineering Practice*, vol. 5, no. 5, pp. 709-719.
- Janusz, M. E. and Venkatasubramanian, V., (1991). Automatic generation of qualitative descriptions of process trends for fault detection and diagnosis, *Engineering Applications of Artificial Intelligence*, vol. 4, pp. 329-339.
- Jia, F., Martin, E. B., and Morris, A. J., (1998). Non-linear Principal Components Analysis for Process Fault Detection, *Computers & Chemical Engineering*, vol. 22, no. 1, p. S851-S854.
- Jutten, C. and Herault, H., (1991). Blind Separation of Sources, Part I: an Adaptive Algorithm Based on Neuromimetic Architecture, *Signal Processing*, vol. 24, pp. 1-10.
- Kano, M., Nagao, K., Hasebe, S., Hashimoto, I., Ohno, H., Strauss, R., and Bakshi, B., (2000). Comparison of statistical process monitoring methods: application to the Eastman challenge problem, *Computers & Chemical Engineering*, vol. 24, pp. 175-181.

- Kano, M., Hasebe, S., Hashimoto, I., and Ohno, H., (2001). A new multivariate statistical process monitoring method using principal component analysis, *Computers & Chemical Engineering*, vol. 25, no. 7-8, pp. 1103-1113.
- Kano, M., Tanaka, S., Ohno, H., Hasebe, S., and Hashimoto, I., (2002a). The Use of Independent Component Analysis for Multivariate Statistical Process Control, Proceedings of International Symposium on Advanced Control of Industrial Processes (AdCONIP'02), Kumamoto, Japan, pp. 423-428.
- Kano, M., Nagao, K., Hasebe, S., Hashimoto, I., Ohno, H., Strauss, R., and Bakshi, B. R., (2002b). Comparison of multivariate statistical process monitoring methods with applications to the Eastman challenge problem, *Computers & Chemical Engineering*, vol. 26, no. 2, pp. 161-174.
- Karpenko, M. and Sepeshri, N., (2002). Neural network classifiers applied to condition monitoring of a pneumatic process valve actuator, *Engineering Applications of Artificial Intelligence*, vol. 15, no. 3-4, pp. 273-283.
- Keller, J.Y., (1999). Fault isolation filter design for linear stochastic systems, *Automatica*, vol. 35, no. 10, pp. 1701-1706.
- Kendra, S. J. and Cinar, A., (1997). Controller performance assessment by frequency domain techniques, *Journal of Process Control*, vol. 7, no. 3, pp. 181-194.
- Kinnaert, M., (1999). Robust fault detection based on observers for bilinear systems, *Automatica*, vol. 35, no. 11, pp. 1829-1842.
- Kinnaert, M., Vrancic, D., Denolin, E., Juricic, D., and Petrovic, J., (2000). Model-based fault detection and isolation for a gas-liquid separation unit, *Control Engineering Practice*, vol. 8, no. 11, pp. 1273-1283.
- Ko, B. S. and Edgar, T. F., (2000). Performance Assessment of Cascade Control Loops, *AIChE Journal*, vol. 46, no. 2, pp. 281-291.
- Ko, B. S. and Edgar, T. F., (2001a). Performance Assessment of Constrained Model Predictive Control Systems, *AIChE Journal*, vol. 47, no. 6, pp. 1363-1371.
- Ko, B. S. and Edgar, T. F., (2001b). Performance assessment of multivariable feedback control systems, *Automatica*, vol. 37, no. 6, pp. 899-905.

- Kozub, D. J., (1997). Controller performance monitoring and diagnosis, Proceedings of Fifth International Conference on Chemical Process Control, CACHE Corp, pp. 83-96.
- Kumamoto, H., Ikenchi, K., Inoue, K., and Henley, E. J., (1984). Application of expert system techniques to fault diagnosis, *The Chemical Engineering Journal*, vol. 29, no. 1, pp. 1-9.
- Kyung, J. M., Gibaek, L., Dong, S. N., Yeo, H. Y., and En, S. Y., (1997). Robust fault diagnosis based on clustered symptom trees, *Control Engineering Practice*, vol. 5, no. 2, pp. 199-208.
- Lee, T. W., (1998). *Independent Component Analysis Theory and applications*, Kluwer Academic Publishers.
- Leger, R. P., Garland, W., and Poehlman, W. F. S., (1998). Fault detection and diagnosis using statistical control charts and artificial neural networks, *Artificial Intelligence in Engineering*, vol. 12, no. 1-2, pp. 35-47.
- Lennox, B., Hiden, H. G., Montague, G. A., Kornfeld, G., and Goulding, P. R., (2000). Application of multivariate statistical process control to batch operations, *Computers & Chemical Engineering*, vol. 24, no. 2-7, pp. 291-296.
- Leonhardt, S. and Ayoubi, M., (1997). Methods of fault diagnosis, *Control Engineering Practice*, vol. 5, no. 5, pp. 683-692.
- Leung, D. and Romagnoli, J., (2000). Dynamic probabilistic model-based expert system for fault diagnosis, *Computers & Chemical Engineering*, vol. 24, no. 11, pp. 2473-2492.
- Li, R. F. and Wang, X. Z., (2002). Dimension reduction of process dynamic trends using independent component analysis, *Computers & Chemical Engineering*, vol. 26, no. 3, pp. 467-473.
- Liu, W., (1999). An extended Kalman filter and neural network cascade fault diagnosis strategy for the glutamic acid fermentation process, *Artificial Intelligence in Engineering*, vol. 13, no. 2, pp. 131-140.
- Ljung, L., (1999). *System Identification Theory for the User*, Second Edition, Prentice Hall PTR.

- Love, P. L. and Simaan, M., (1988). Automatic recognition of primitive changes in manufacturing process signals, *Pattern Recognition*, vol. 21, pp. 333-342.
- Lynch, C. B. and Dumont, G. A., (1996). Control loop performance monitoring, *IEEE Trans. Control Sys. Tech.*, vol. 4, no. 2, pp. 185-192.
- MacGregor, J. F. and Kourti, T., (1995). Statistical process control of multivariate processes, *Control Engineering Practice*, vol. 3, no. 3, pp. 403-414.
- Mangoubi, R. S., (1998). *Robust estimation and failure detection*, Springer.
- Martin, G. D., Turpin, L. E., and Cline, R. P., (1991). Estimating control function benefits, *Hydrocarbon Processing* (June), pp. 68-73.
- Martin, E. B., Morris, A. J., and Zhang, J., (1996). Process performance monitoring using multivariate statistical process control, *Iee Proceedings-Control Theory and Applications*, vol. 143, no. 2, pp. 132-144.
- McAvoy, T. J. and Ye, N., (1994). Base Control for the Tennessee Eastman Problem, *Computers and Chemical Engineering*, vol. 18, no. 5, pp. 383-413.
- McAvoy, T. J., Ye, N., and Chen, G., (1995). An Improved Base Control for the Tennessee Eastman Problem, *Proceedings of American Control Conference*, Seattle, Washington, pp. 240-244.
- McGhee, J., Henderson, I. A., and Baird, A., (1997). Neural networks applied for the identification and fault diagnosis of process valves and actuators, *Measurement*, vol. 20, no. 4, pp. 267-275.
- McKenzie, F. D., Gonzalez, A. J., and Morris, R., (1998). An integrated model-based approach for real-time on-line diagnosis of complex systems, *Engineering Applications of Artificial Intelligence*, vol. 11, no. 2, pp. 279-291.
- Medvedev, A., (1995). Fault Detection and Isolation by a Continuous Parity Space Method, *Automatica*, vol. 31, no. 7, pp. 1039-1044.
- Miao, T. and Seborg, D., (1999). Automatic detection of excessively oscillatory feedback control loops, *IEEE Int. Conf. On Control Applications*, Hawaii, USA.
- Mohindra, S. and Clark, P. A., (1993). A distributed fault diagnosis method based on digraph models: Steady-state analysis, *Computers & Chemical Engineering*, vol. 17, no. 2, pp. 193-209.

- Montmain, J. and Gentil, S., (2000). Dynamic causal model diagnostic reasoning for online technical process supervision, *Automatica*, vol. 36, no. 8, pp. 1137-1152.
- Mylaraswamy, D. and Venkatasubramanian, V., (1997). A Hybrid Framework for Large Scale Process Fault Diagnosis, *Computers & Chemical Engineering*, vol. 21, no. 1, p. S935-S940.
- Ouali, M.-S., Ait-Kadi, D., and Rezg, N., (1999). Fault diagnosis model based on Petri net with fuzzy colors, *Computers & Industrial Engineering*, vol. 37, no. 1-2, pp. 173-176.
- Owen, J. G., Read, D., Blekkenhorst, H., and Roche, A., (1996). A mill prototype for automatic monitoring of control loop performance, *Proceedings Control Systems 96*, pp. 171-178.
- Ozyurt, B., Sunol, A. K., Camurdan, M. C., Mogili, P., and Hall, L. O., (1998). Chemical plant fault diagnosis through a hybrid symbolic-connectionist machine learning approach, *Computers & Chemical Engineering*, vol. 22, no. 1-2, pp. 299-321.
- Papoulis, A., (1991). *Probability, Random Variables, and Stochastic Process*, Third Edition, McGraw-Hill, Inc..
- Patton, R. J., Frank, P. M., and Clark, R. N., (1989). *Fault diagnosis in dynamic systems: theory and applications*, Prentice Hall.
- Patton, R. J. and Chen, J., (1997). Observer-based fault detection and isolation: robustness and applications, *Control Engineering Practice*, vol. 5, no. 5, pp. 671-682.
- Patton, R. J. and Hou, M., (1998). Design of Fault Detection and Isolation Observers: A Matrix Pencil Approach, *Automatica*, vol. 34, no. 9, pp. 1135-1140.
- Patton, R. J., Frank, P. M., and Clark, R. N., (2000). *Issues of fault diagnosis for dynamic systems*, Springer.
- Paulonis, M. A. and Cox, J. W., (2003). A practical approach for large-scale controller performance assessment, diagnosis, and improvement, *Journal of Process Control*, vol. 13, no. 2, pp. 155-168.

- Pranatyasto, T. N. and Qin, S. J., (2001). Sensor validation and process fault diagnosis for FCC units under MPC feedback, *Control Engineering Practice*, vol. 9, no. 8, pp. 877-888.
- Qin, S. J., (1998). Control performance monitoring -- a review and assessment, *Computers & Chemical Engineering*, vol. 23, no. 2, pp. 173-186.
- Raich, A. and Cinar, A., (1997). Diagnosis of process disturbances by statistical distance and angle measures, *Computers & Chemical Engineering*, vol. 21, no. 6, pp. 661-673.
- Rakar, A., Juricic, D., and Balle, P., (1999). Transferable belief model in fault diagnosis, *Engineering Applications of Artificial Intelligence*, vol. 12, no. 5, pp. 555-567.
- Rengaswamy, R. and Venkatasubramanian, V., (1995). A Syntactic Pattern-recognition Approach for Process Monitoring and Fault Diagnosis, *Engineering Applications of Artificial Intelligence*, vol. 8, no. 1, pp. 35-51.
- Rengaswamy, R. and Venkatasubramanian, V., (2000). A fast training neural network and its updation for incipient fault detection and diagnosis, *Computers & Chemical Engineering*, vol. 24, no. 2-7, pp. 431-437.
- Rengaswamy, R., Hagglund, T., and Venkatasubramanian, V., (2001a). A qualitative shape analysis formalism for monitoring control loop performance, *Engineering Applications of Artificial Intelligence*, vol. 14, no. 1, pp. 23-33.
- Rengaswamy, R., Mylaraswamy, D., Arzen, K.-E., and Venkatasubramanian, V., (2001b). A comparison of model-based and neural network-based diagnostic methods, *Engineering Applications of Artificial Intelligence*, vol. 14, no. 6, pp. 805-818.
- Rich, S. H. and Venkatasubramanian, V., (1987). Model-based reasoning in diagnostic expert systems for chemical process plants, *Computers & Chemical Engineering*, vol. 11, no. 2, pp. 111-122.
- Rozier, D., (2001). A strategy for diagnosing complex multiple-fault situations with a higher accuracy/cost ratio, *Engineering Applications of Artificial Intelligence*, vol. 14, no. 2, pp. 217-227.

- Ruiz, D., Nougues, J., Calderon, Z., Espuna, A., and Puigjaner, L., (2000). Neural network based framework for fault diagnosis in batch chemical plants, *Computers & Chemical Engineering*, vol. 24, no. 2-7, pp. 777-784.
- Ruiz, D., Canton, J., Maria Nougues, J., Espuna, A., and Puigjaner, L., (2001a). On-line fault diagnosis system support for reactive scheduling in multipurpose batch chemical plants, *Computers & Chemical Engineering*, vol. 25, no. 4-6, pp. 829-837.
- Ruiz, D., Nougues, J. M., and Puigjaner, L., (2001b). Fault diagnosis support system for complex chemical plants, *Computers & Chemical Engineering*, vol. 25, no. 1, pp. 151-160.
- Russell, E. L., Chiang, L. H., and Braatz, R. D., (2000). *Data-driven techniques for fault detection and diagnosis in chemical processes*, Springer.
- Sanchez A, V. D., (2002). Frontiers of research in BSS/ICA, *Neurocomputing*, vol. 49, no. 1-4, pp. 7-23.
- Schiller, F., Schroder, J., and Lunze, J., (2001). Diagnosis of transient faults in quantised systems, *Engineering Applications of Artificial Intelligence*, vol. 14, no. 4, pp. 519-536.
- Shao, R., Jia, F., Martin, E. B., and Morris, A. J., (1999). Wavelets and non-linear principal components analysis for process monitoring, *Control Engineering Practice*, vol. 7, no. 7, pp. 865-879.
- Shen, L. C. and Hsu, P. L., (1998). Robust Design of Fault Isolation Observers, *Automatica*, vol. 34, no. 11, pp. 1421-1429.
- Shinskey, F. G., (1990). How good are our controllers in absolute performance and robustness, *Measurement and control*, vol. 23, no. May, pp. 114-121.
- Shum, S. K., Davis, J. F., Punch, I. I. I., and Chandrasekaran, B., (1988). An expert system approach to malfunction diagnosis in chemical plants, *Computers & Chemical Engineering*, vol. 12, no. 1, pp. 27-36.
- Simani, S., Fantuzzi, C., and Beghelli, S., (2000). Diagnosis techniques for sensor faults of industrial processes, *Ieee Transactions on Control Systems Technology*, vol. 8, no. 5, pp. 848-855.

- Simani, S., Fantuzzi, C., and Patton, R. J., (2003). *Model-based fault diagnosis in dynamic systems using identification techniques*, Springer.
- Simoglou, A., Martin, E. B., and Morris, A. J., (2002). Statistical performance monitoring of dynamic multivariate processes using state space modelling, *Computers & Chemical Engineering*, vol. 26, no. 6, pp. 909-920.
- Stanfelj, N. A., Marlin, T. E., and MacGregor, J. F., (1993). Monitoring and diagnosis process control performance: the single-loop case, *Ind.Eng.Chem.Res.*, vol. 32, pp. 301-314.
- Stephanopoulos, G., (1984). *Chemical process control: an introduction to theory and practice*, Prentice Hall International.
- Sun, X., Chen, T., and Marquez, H. J., (2002). Efficient model-based leak detection in boiler steam-water systems, *Computers & Chemical Engineering*, vol. 26, no. 11, pp. 1643-1647.
- Taha, O., Dumont, G. A., and Davies, M. S., (1996). Detection and diagnosis of oscillations in control loops, Proceedings of the 35th conference on decision and control, Kobe, Japan, pp. 2432-2437.
- Tan, C. P. and Edwards, C., (2002). Sliding mode observers for detection and reconstruction of sensor faults, *Automatica*, vol. 38, no. 10, pp. 1815-1821.
- Tarifa, E. E. and Scenna, N. J., (1997). Fault diagnosis, direct graphs, and fuzzy logic, *Computers & Chemical Engineering*, vol. 21, no. 1, p. S649-S654.
- Teo, C. Y. and Gooi, H. B., (1997). Application of Knowledge-Based Systems for Fault Diagnosis and Supply Restoration, *Engineering Applications of Artificial Intelligence*, vol. 10, no. 6, pp. 631-638.
- Thornhill, N. F. and Hägglund, T., (1997). Detection and diagnosis of oscillation in control loops, *Control Engineering Practice*, vol. 5, no. 10, pp. 1343-1354.
- Thornhill, N. F., Oettinger, M., and Fedenczuk, P., (1999). Refinery-wide control loop performance assessment, *Journal of Process Control*, vol. 9, no. 2, pp. 109-124.
- Thornhill, N. F., Shah, S. L., and Huang, B., (2000). Detection and diagnosis of unit-wide oscillations, Process Control and Instrumentation 2000, Glasgow, Scotland.

- Thornhill, N. F., Shah, S. L., and Huang, B., (2001). Detection of distributed oscillations and root cause diagnosis, 4TH IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries, Korea, pp. 167-172.
- Thornhill, N. F., Shah, S. L., Huang, B., and Vishnubhotla, A., (2002a). Spectral principal component analysis of dynamic process data, *Control Engineering Practice*, vol. 10, no. 8, pp. 833-846.
- Thornhill, N. F., Xia, C., Howell, J., Cox, J., and Paulonis, M., (2002b). Analysis of plant-wide disturbances through data-driven techniques and process understanding, 15TH IFAC World Congress, Barcelona.
- Thornhill, N. F., Huang, B., and Zhang, H., (2003). Detection of multiple oscillations in control loops, *Journal of Process Control*, vol. 13, no. 1, pp. 91-100.
- Tse, P., Wang, D., and Atherton, D., (1996). Harmony Theory Yields Robust Machine Fault-Diagnostic Systems Based on Learning Vector Quantization Classifiers, *Engineering Applications of Artificial Intelligence*, vol. 9, no. 5, pp. 487-498.
- Tsuge, Y., Hiratsuka, K., Takeda, K., and Matsuyama, H., (2000). A fault detection and diagnosis for the continuous process with load-fluctuations using orthogonal wavelets, *Computers & Chemical Engineering*, vol. 24, no. 2-7, pp. 761-767.
- Tyler, M. L. and Morari, M., (1996). Performance Monitoring of Control Systems using Likelihood Methods, *Automatica*, vol. 32, no. 8, pp. 1145-1162.
- Vedam, H. and Venkatasubramanian, V., (1997). Signed Digraph Based Multiple Fault Diagnosis, *Computers & Chemical Engineering*, vol. 21, no. 1, p. S655-S660.
- Vedam, H. and Venkatasubramanian, V., (1999). PCA-SDG based process monitoring and fault diagnosis, *Control Engineering Practice*, vol. 7, no. 7, pp. 903-917.
- Verde, C., (2001). Multi-leak detection and isolation in fluid pipelines, *Control Engineering Practice*, vol. 9, no. 6, pp. 673-682.
- Vora, N., Tambe, S. S., and Kulkarni, B. D., (1997). Counterpropagation neural networks for fault detection and diagnosis, *Computers & Chemical Engineering*, vol. 21, no. 2, pp. 177-185.
- Wang, X. Z., Yang, S. A., Veloso, E., Lu, M. L., and McGreavy, C., (1995). Qualitative process modelling - a fuzzy signed directed graph method, *Computers & Chemical Engineering*, vol. 19, no. 1, pp. 735-740.

- Weerasinghe, M., Gomm, J. B., and Williams, D., (1998). Neural networks for fault diagnosis of a nuclear fuel processing plant at different operating points, *Control Engineering Practice*, vol. 6, no. 2, pp. 281-289.
- Welch, P. D., (1967). The use of fast Fourier transforms for the estimation of power spectra, *IEEE Transactions on Audio and Electroacoustics*, vol. AU-15, pp. 70-73.
- Weyer, E., Szederkenyi, G., and Hantos, K., (2000). Grey box fault detection of heat exchangers, *Control Engineering Practice*, vol. 8, no. 2, pp. 121-131.
- Wilcox, N. A. and Himmelblau, D. M., (1994). The possible cause and effect graphs (PCEG) model for fault diagnosis--I. Methodology, *Computers & Chemical Engineering*, vol. 18, no. 2, pp. 103-116.
- Wong, J. C., McDonald, K. A., and Palazoglu, A., (2001). Classification of abnormal plant operation using multiple process variable trends, *Journal of Process Control*, vol. 11, no. 4, pp. 409-418.
- Xia, C. and Howell, J., (2001a). Controller output based, single number statistics for control loop status monitoring, 4TH IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries, Korea, pp. 127-132.
- Xia, C. and Howell, J., (2001b). Loop status statistics, 4TH IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries, Korea, pp. 372-376.
- Xia, C. and Howell, J., (2002a). PI loop status monitoring, IEEE International Conference on Control Applications, Glasgow.
- Xia, C. and Howell, J., (2002b). Fault localization based on loop status monitoring, IEEE International Conference on Control Applications, Glasgow.
- Xia, C., Howell, J., and Thornhill, N. F., (2003). Detecting and isolating multiple plant-wide oscillations via spectral independent component analysis, *Submitted to Automatica*.
- Xia, C. and Howell, J., (2003a). Isolating multiple sources of plant-wide oscillations via independent component analysis, Safeprocess 2003, Washington, D.C., USA.
- Xia, C. and Howell, J., (2003b). Loop status monitoring and fault localisation, *Journal of Process Control (In press)*.

- Yang, S. H., Chen, B. H., and Wang, X. Z., (2000). Neural network based fault diagnosis using unmeasurable inputs, *Engineering Applications of Artificial Intelligence*, vol. 13, no. 3, pp. 345-356.
- Yoon, S. and MacGregor, J. F., (2001). Fault diagnosis with multivariate statistical models part I: using steady state fault signatures, *Journal of Process Control*, vol. 11, no. 4, pp. 387-400.
- Yu, D. and Shields, D. N., (1996). A Bilinear Fault Detection Observer, *Automatica*, vol. 32, no. 11, pp. 1597-1602.
- Yu, D. L., Gomm, J. B., and Williams, D., (1999). Sensor fault diagnosis in a chemical process via RBF neural networks, *Control Engineering Practice*, vol. 7, no. 1, pp. 49-55.
- Zhang, Q., Basseville, M., and Benveniste, A., (1998). Fault Detection and Isolation in Nonlinear Dynamic Systems: A Combined Input-Output and Local Approach, *Automatica*, vol. 34, no. 11, pp. 1359-1373.
- Zhao, J., Chen, B., and Shen, J., (1997). A hybrid ANN-ES system for dynamic fault diagnosis of hydrocracking process, *Computers & Chemical Engineering*, vol. 21, no. 1, p. S929-S933.
- Zhao, J., Chen, B., and Shen, J., (1998). Multidimensional non-orthogonal wavelet-sigmoid basis function neural network for dynamic process fault diagnosis, *Computers & Chemical Engineering*, vol. 23, no. 1, pp. 83-92.
- Zolghadri, A., Henry, D., and Monsion, M., (1996). Design of nonlinear observers for fault diagnosis: a case study, *Control Engineering Practice*, vol. 4, no. 11, pp. 1535-1544.

Appendix A

Proofs for the Basis in Chapter 3

A.1 Proof of Theorem 3-1

Assume that a closed loop has a sine wave oscillatory trend with frequency ω , then:

$$\hat{y}(t) = A \sin(\omega t) \quad (\text{A-1})$$

With reference to Section 3.2.1, assume that the power spectrum, $G_{e_y}(\omega)$, of e_y is uniform over the bandwidth, that is,

$$G_{e_y}(\omega) = \begin{cases} G & B_0 \leq \omega \leq B \\ 0 & \begin{cases} \omega > B \\ \omega < B_0 \end{cases} \end{cases} \quad (\text{A-2})$$

where *Nyquist* angular frequency $B = \pi/T_s$ (T_s is sampling time interval) and B_0 is low-band limit. Define $K_B: B_0 = B/K_B$.

Variance of e_y : $\sigma_{e_y}^2$

$$\sigma_{e_y}^2 = R_{e_y}(0) = \int_{B_0}^B G_{e_y}(\omega) d\omega = G(B - B_0) \quad (\text{A-3})$$

Variance of e_u : $\sigma_{e_u}^2$

According to the input/output auto-spectrum relationship (Bendat & Piersol 1993), the auto-spectrum of e_u can be given by:

$$G_{e_u}(\omega) = |H(j\omega)|^2 G_{e_y}(\omega) = K_p^2 \left(1 + \frac{1}{T_i^2 \omega^2} \right) G_{e_y}(\omega) \quad (\text{A-4})$$

where K_p and T_i are PI controller parameters, and the variance of e_u is:

$$\sigma_{e_u}^2 = \int_{B_0}^B G_{e_u}(\omega) d\omega = G(B - B_0) K_p^2 Q \quad (\text{A-5})$$

where

$$Q = 1 + \frac{K_B}{\pi^2} \left[\frac{T_s}{T_i} \right]^2 \quad (\text{A-6})$$

Variations of \hat{y} and \hat{u} : $\sigma_{\hat{y}}^2, \sigma_{\hat{u}}^2$

The s -domain expression of $\hat{y}(t)$ is:

$$\hat{y}(s) = \frac{A\omega}{s^2 + \omega^2} \quad (\text{A-7})$$

and

$$\hat{u}(s) = H(s)\hat{y}(s) = K_p \left[1 + \frac{1}{T_i s} \right] \left(\frac{A\omega}{s^2 + \omega^2} \right) \quad (\text{A-8})$$

The controller output can be obtained by Laplace inverse transformation, as

$$K_p \left[A \sin \omega t + \frac{A}{T_i \omega} (1 - \cos \omega t) \right] \quad (\text{A-9})$$

Since the data analysis is based on deviation variables, form $\hat{u}(t)$ by removing the

constant term $\frac{K_p A}{T_i \omega}$ to give

$$\hat{u}(t) = K_p A \sqrt{M} \sin(\omega t + \beta) \quad (\text{A-10})$$

where $M = 1 + \frac{1}{(T_i \omega)^2}$, and $\beta = -\tan^{-1}\left(\frac{1}{T_i \omega}\right)$.

The variance of deviations $\hat{u}(t)$ and $\hat{y}(t)$ can be given by:

$$\sigma_{\hat{u}}^2 = \frac{1}{2} K_p^2 A^2 M \quad (\text{A-11})$$

$$\sigma_{\hat{y}}^2 = \frac{1}{2} A^2 \quad (\text{A-12})$$

Signal-to-noise ratio indexes: η_y and η_u

Substituting the above equations for $\sigma_{\hat{y}}^2, \sigma_{e_y}^2, \sigma_{\hat{u}}^2$ and $\sigma_{e_u}^2$ into Equations (3.4) & (3.5) in Definition 3-1 yields

$$\eta_y = \frac{\frac{1}{2} A^2}{G(B-B_0)} \quad (\text{A-13})$$

$$\eta_u = \frac{\frac{1}{2} A^2}{G(B-B_0)} \frac{1 + \frac{1}{(T_i \omega)^2}}{Q} = \eta_y \underbrace{\left(\frac{1 + \frac{1}{(T_i \omega)^2}}{Q} \right)}_{J(\omega)} = J(\omega) \eta_y \quad (\text{A-14})$$

Ratio of indexes: R

Substituting the above equations for η_y and η_u into Equation (3.6) yields

$$R = \frac{\eta_y}{\eta_u} = \frac{\eta_y}{J(\omega) \eta_y} = \frac{1}{J(\omega)} = \frac{Q}{1 + \frac{1}{(T_i \omega)^2}} \quad (\text{A-15})$$

Normalised ratio R_n can be given by

$$R_n = \frac{R}{Q} \in (0,1) \quad (\text{A-16})$$

A.2 General Deterministic Oscillatory Trends (PI)

A more general expression for the deterministic oscillatory trend $\hat{y}(t)$ is a Fourier series representation of sine waves with n components,

$$\hat{y}(t) = \left(\sum_{k=1}^n A_k \sin(k\omega t) \right), \quad (\text{A-17})$$

where component $A_1 \sin(\omega t)$ is the dominant fundamental component, i.e., ω is the fundamental frequency and $A_1 \gg A_k, k = 2 \dots n$.

Applying the Principle of Linear Superposition, the controller output deviation can be derived as follows:

$$\hat{u}(t) = K_p \sum_{k=1}^n A_k \sqrt{M_k} \sin(k\omega t + \beta_k), \quad (\text{A-18})$$

where $M_k = 1 + \frac{1}{(T_i k \omega)^2}$ and $\beta_k = -\tan^{-1}\left(\frac{1}{T_i k \omega}\right)$

The variance of $\hat{u}(t)$ is:

$$\begin{aligned} \sigma_{\hat{u}}^2 &= E(\hat{u}^2) = K_p^2 E \left[\left(\sum_{k=1}^n A_k \sqrt{M_k} \sin(k\omega t + \beta_k) \right)^2 \right] \\ &= K_p^2 \left[\frac{1}{2} \sum_{k=1}^n A_k^2 M_k + \underbrace{\sum_{i,j} A_i A_j \sqrt{M_i M_j} E(\sin(i\omega t + \beta_i) \sin(j\omega t + \beta_j))}_{ERR} \right] \quad (\text{A-19}) \\ &= K_p^2 \left(\frac{1}{2} \sum_{k=1}^n A_k^2 M_k + ERR \right) = K_p^2 \left(\frac{1}{2} \sum_{k=1}^n A_k^2 M_k \right), \end{aligned}$$

where $i, j \in [1, n], i \neq j$. $ERR=0$, because $\sin(i\omega t + \beta_i)$ and $\sin(j\omega t + \beta_j)$ are orthogonal.

The variance of $\hat{y}(t)$ is:

$$\sigma_{\hat{y}}^2 = \frac{1}{2} \sum_{k=1}^n A_k^2 \quad (\text{A-20})$$

The ratio R can be given by

$$R = \frac{\eta_y}{\eta_u} = \frac{\frac{1}{2} \sum_{k=1}^n A_k^2}{GB} = \frac{Q \sum_{k=1}^n A_k^2}{K_p^2 \left(\frac{1}{2} \sum_{k=1}^n A_k^2 M_k \right)} = \frac{Q \sum_{k=1}^n A_k^2}{\sum_{k=1}^n A_k^2 M_k} \quad (\text{A-21})$$

$$= \frac{Q \sum_{k=1}^n A_k^2}{GB K_p^2 Q}$$

Due to the domination feature of component 1 in the trend, i.e., $A_1 \gg A_k, k = 2 \dots n$, Normalised R_n can be approximately given by:

$$R_n = \frac{\sum_{k=1}^n A_k^2}{\sum_{k=1}^n A_k^2 M_k} \approx \frac{A_1^2}{A_1^2 M_1} = \frac{1}{M_1} = \frac{1}{1 + \frac{1}{(T_i \omega)^2}}, \quad (\text{A-22})$$

which is the same as the one derived from the single sine-wave trend case.

The determination threshold, $R_n \geq 0.8$, when $\omega \geq \frac{2}{T_i}$, can also be used to determine

whether there is a high fundamental-frequency periodical trend in a closed loop.

A.3 Oscillatory Trend with Frequency ω (PID)

For a generalised PID controller the ratios R and R_n can be represented as:

$$R = \frac{Q_D}{M_D}, \quad (\text{A-23})$$

$$R_n = \frac{R}{Q_D} = \frac{1}{M_D}, \quad (\text{A-24})$$

where Q_D is a constant and M_D is a frequency dependent function. If the controller has a transfer function $H(s)$, the analysis of Appendix A.1 can be repeated to produce the following expression for Q_D :

$$Q_D = \frac{1}{K_p^2} \frac{\int_{B_0}^B |H(j\omega)|^2 d\omega}{B - B_0} \quad (\text{A-25})$$

$$M_D = \frac{1}{K_p^2} |H(j\omega)|^2 \quad (\text{A-26})$$

where K_p is the controller gain, $B = \frac{\pi}{T_s}$, and B_0 is the low band limit and can be

chosen as the notch frequency $\frac{1}{\sqrt{T_i T_d}}$. Thus for a PID controller with transfer

function

$$K_p \left(1 + \frac{1}{T_i s} + T_d s \right)$$

where T_d is the derivative time constant,

$$Q_D = \left(1 - \frac{2}{K_d} \right) + \frac{\pi^2}{3K_d^2} \left(\frac{T_i}{T_s} \right)^2 + \frac{K_B}{\pi^2} \left(\frac{T_s}{T_i} \right)^2, \quad (\text{A-27})$$

$$M_D = \left(1 - \frac{2}{K_d}\right) + \frac{(T_i\omega)^2}{K_d^2} + \frac{1}{(T_i\omega)^2} \quad (\text{A-28})$$

and K_d is the ratio of T_i/T_d .

When $K_d \rightarrow \infty$, i.e., there is no derivative action, R will equal Q/M , i.e. the result will be the same as that for PI control. When K_d is a normal constant value, the $R_n - \omega$ relationship will differ from that of *PI* control.

Appendix B

Analysis Results of Real Data Set From an Eastman

Chemical Company Plant

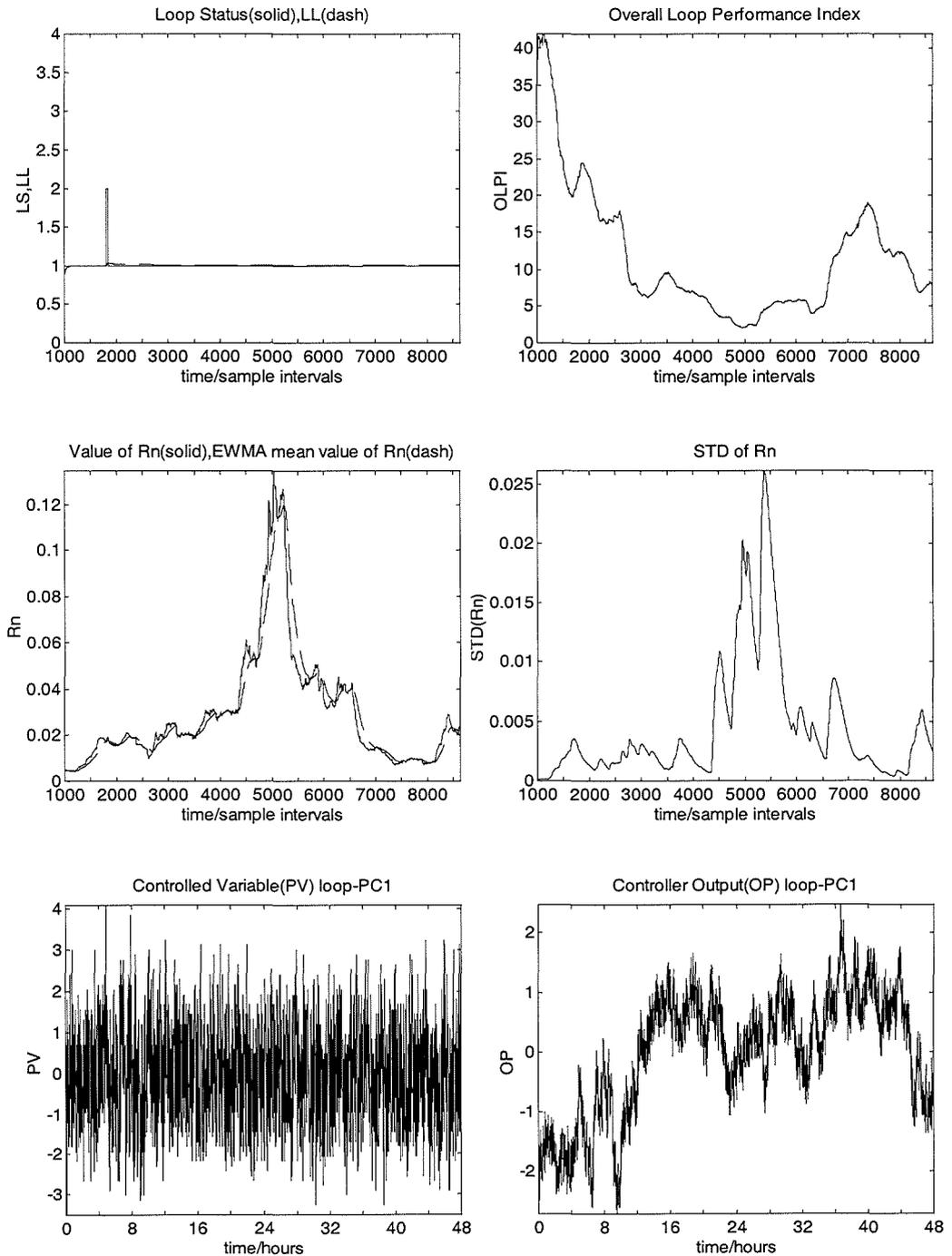


Figure B-1: Results for loop PC1

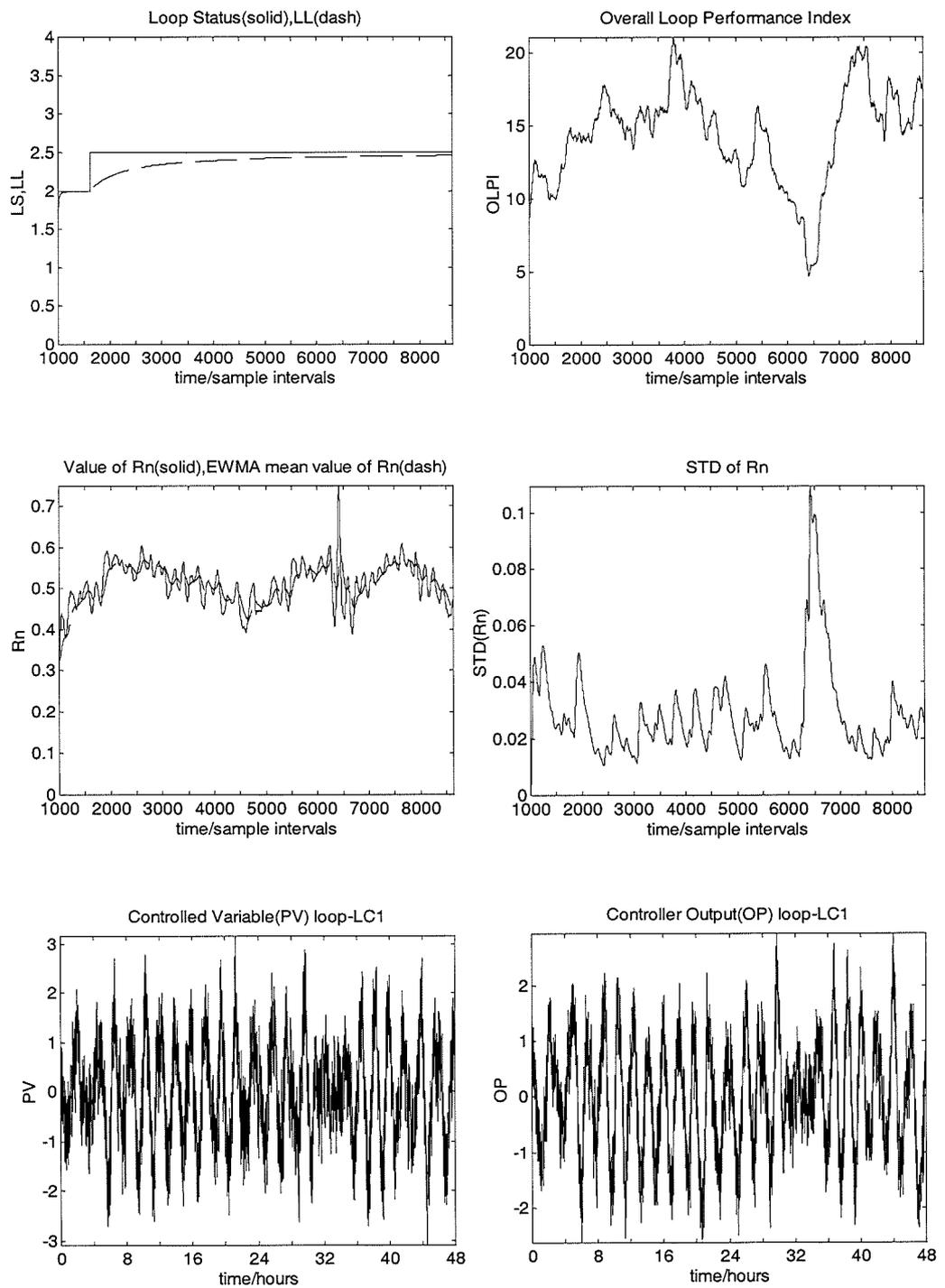


Figure B-2: Results for loop LC1

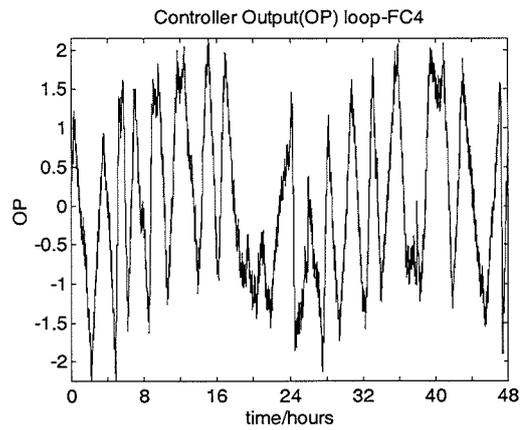
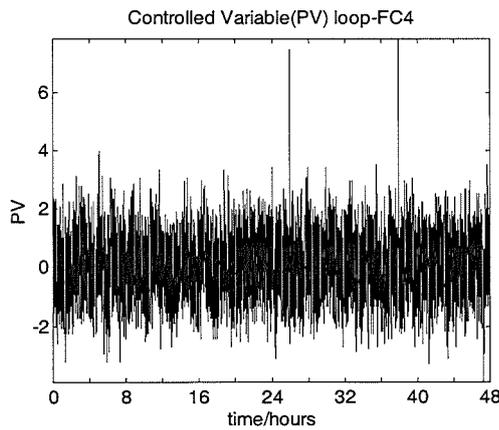
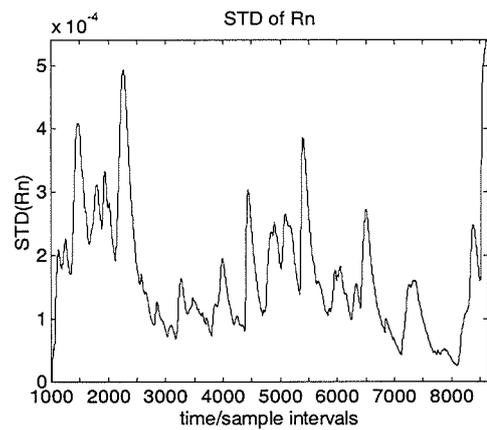
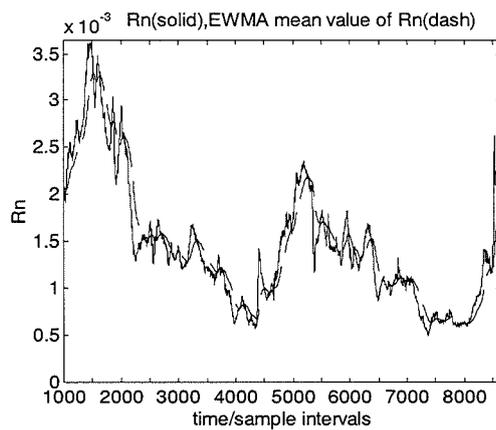
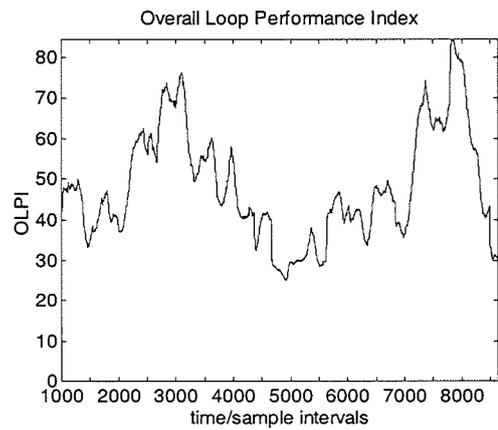
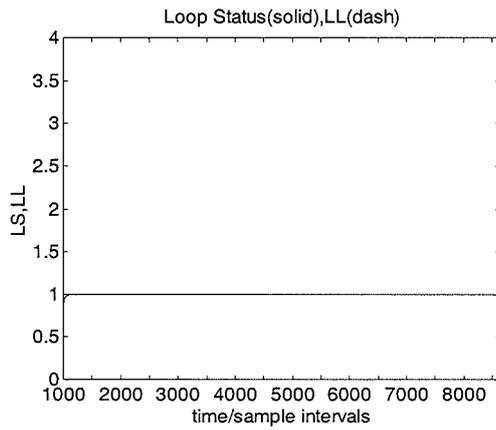


Figure B-3: Results for loop FC4

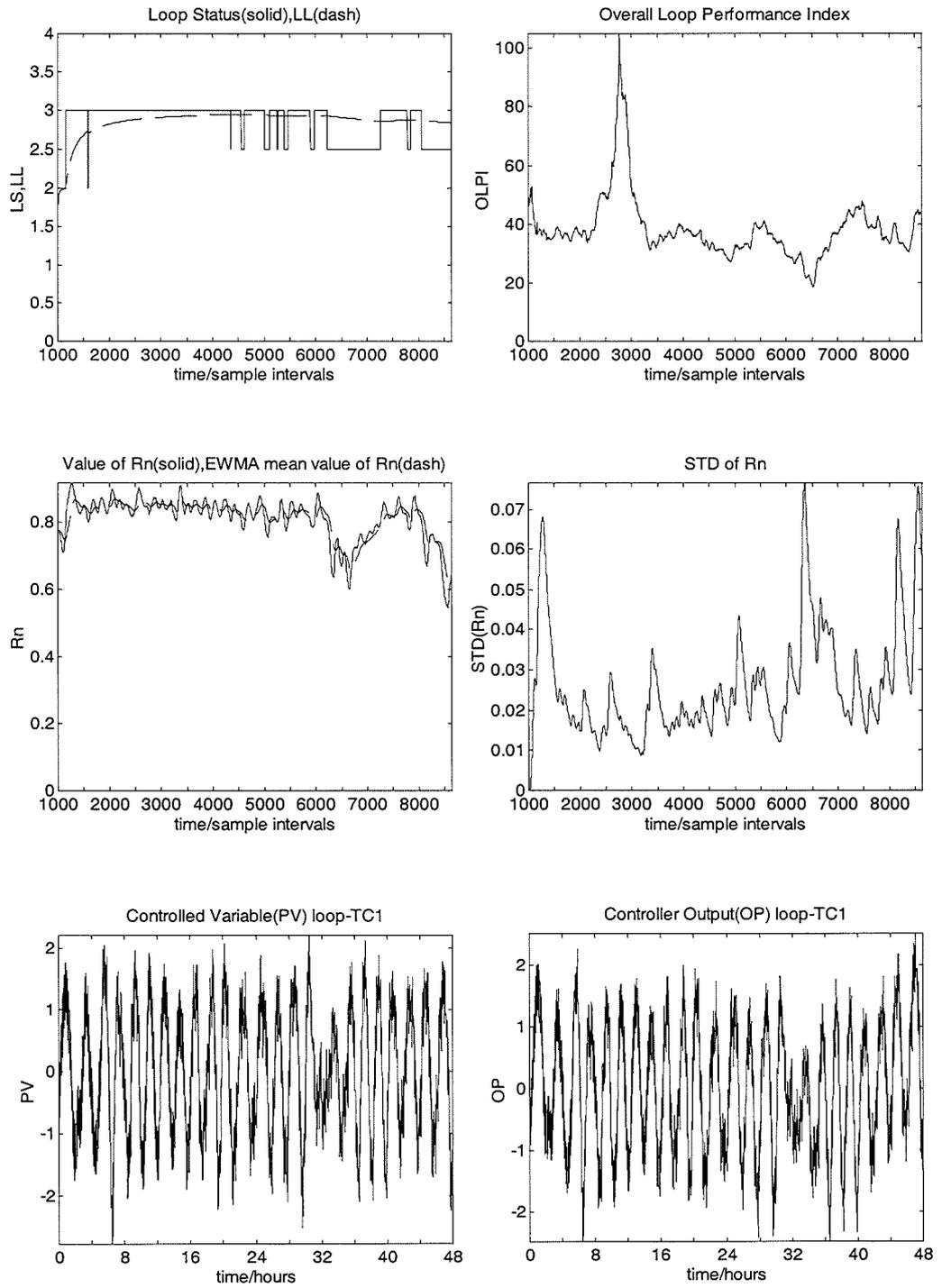


Figure B-4: Results for loop TC1

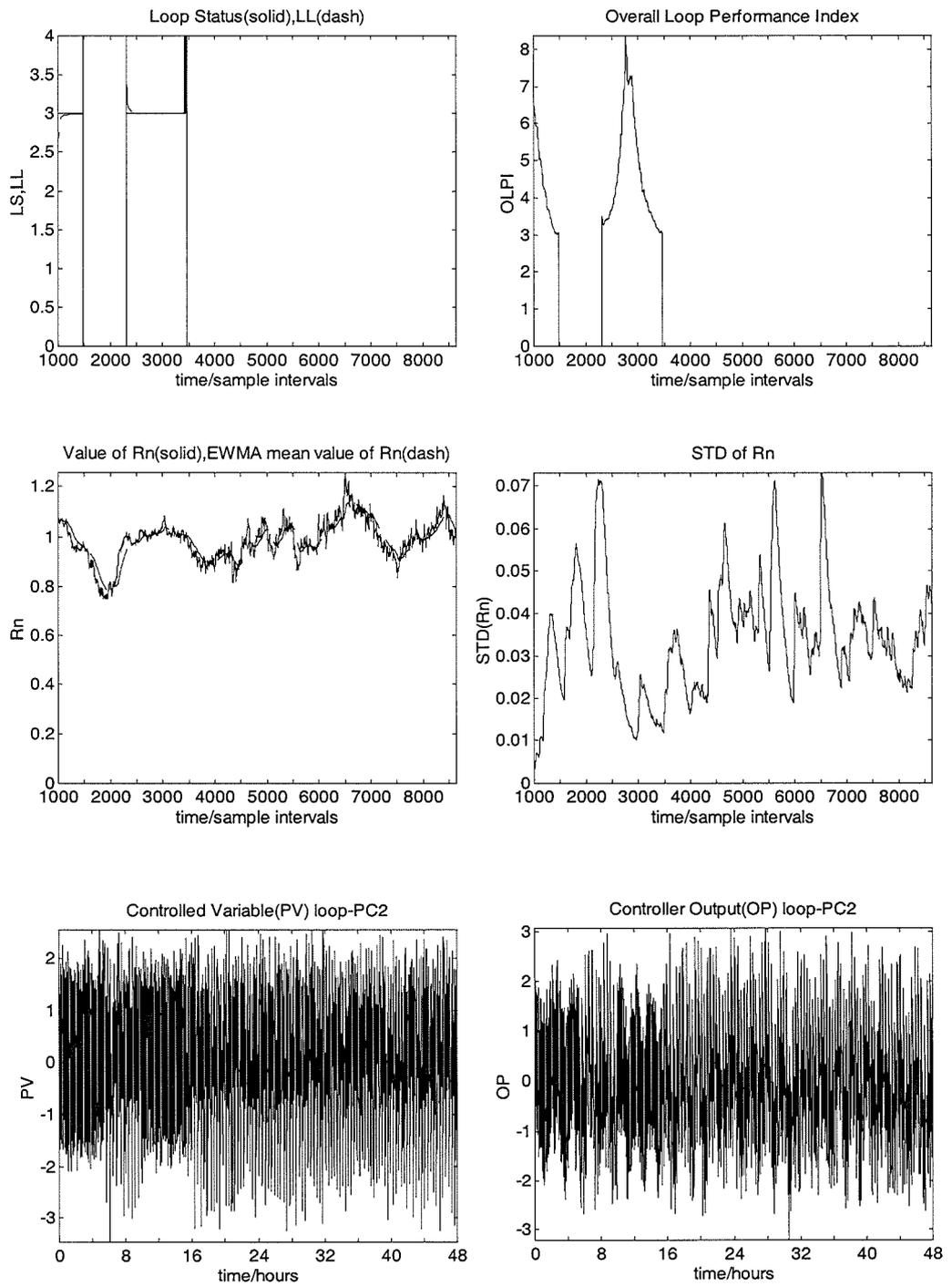


Figure B-5: Results for loop PC2

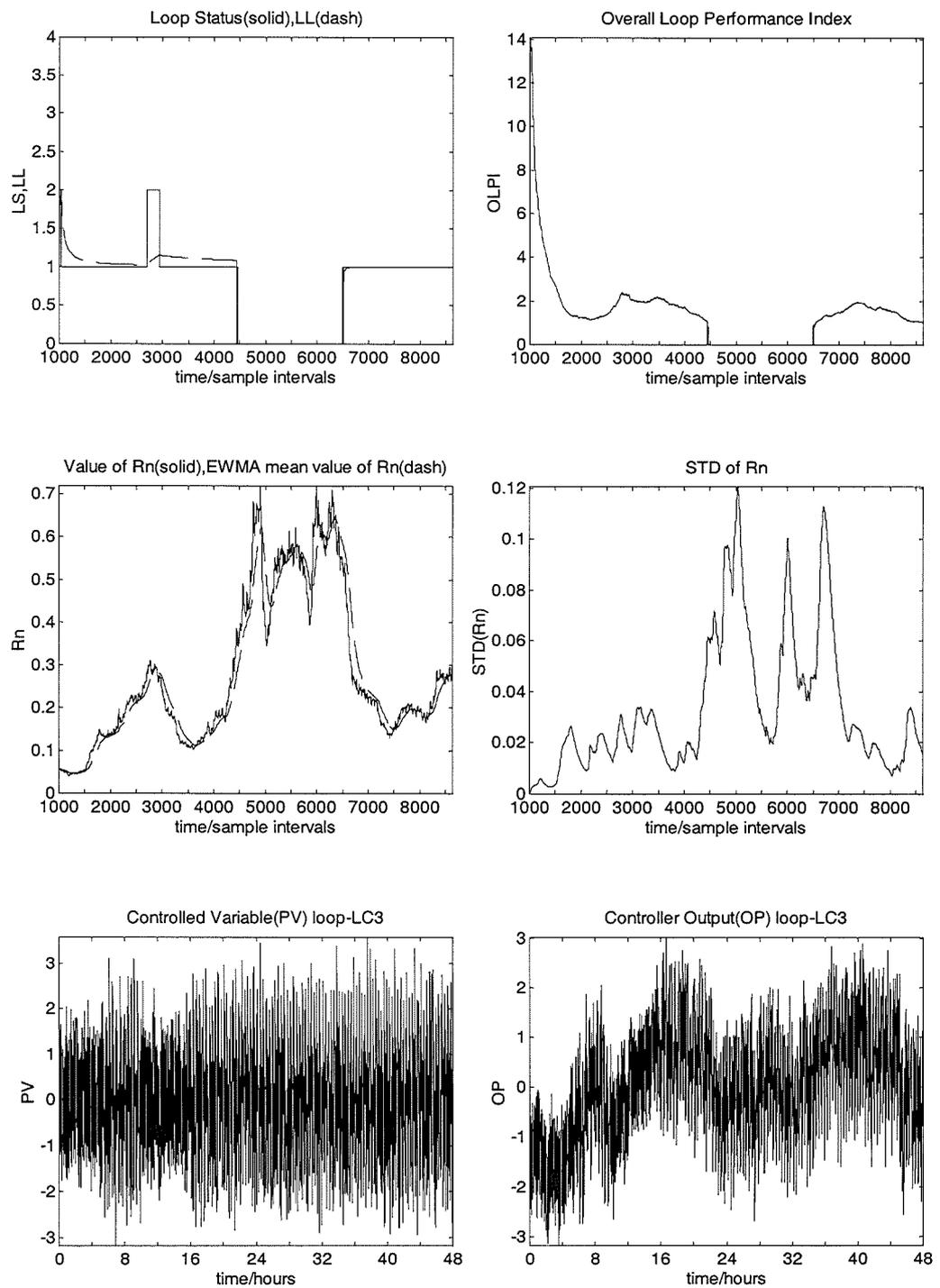


Figure B-6: Results for loop LC3

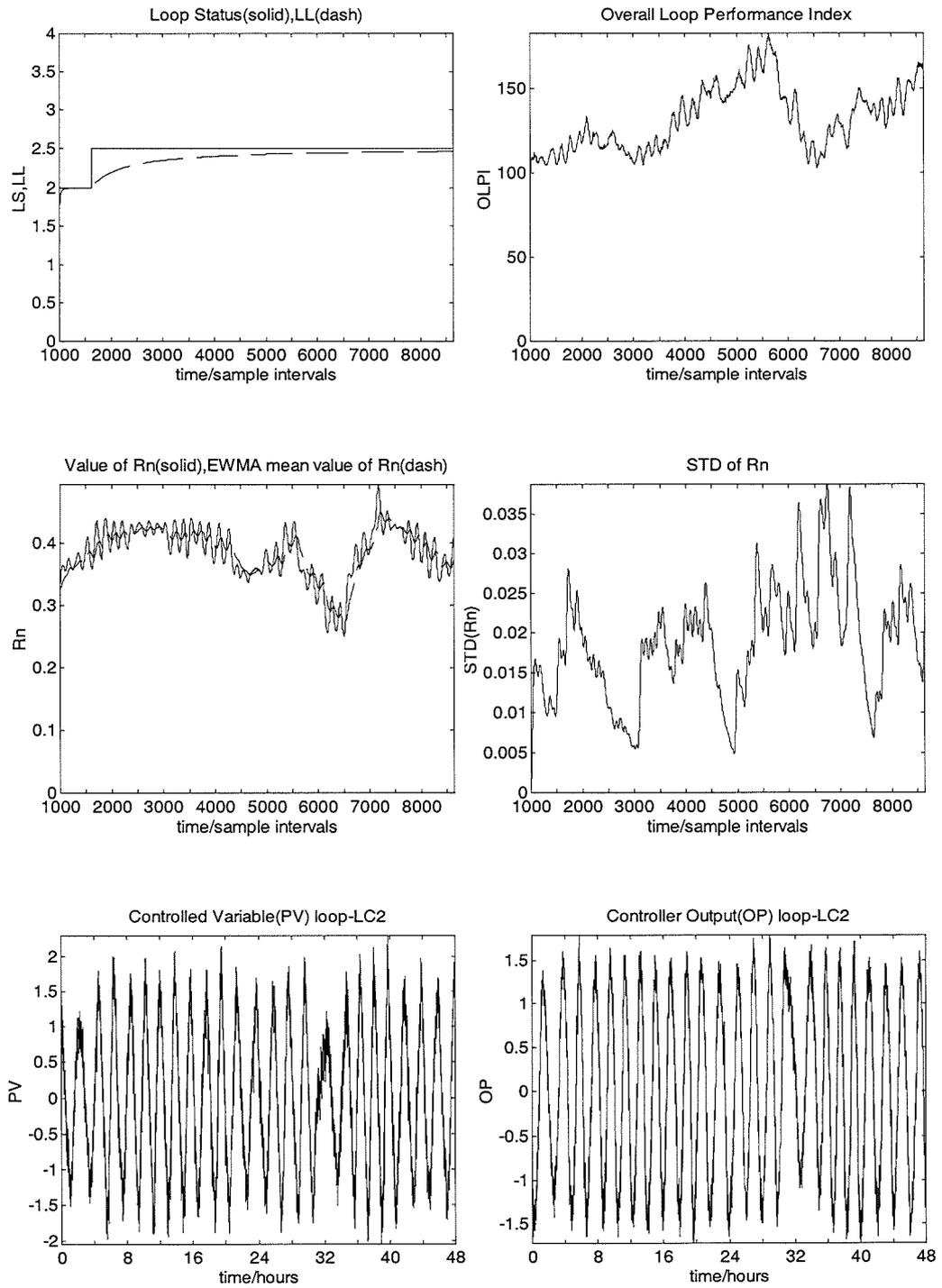


Figure B-7: Results for loop LC2

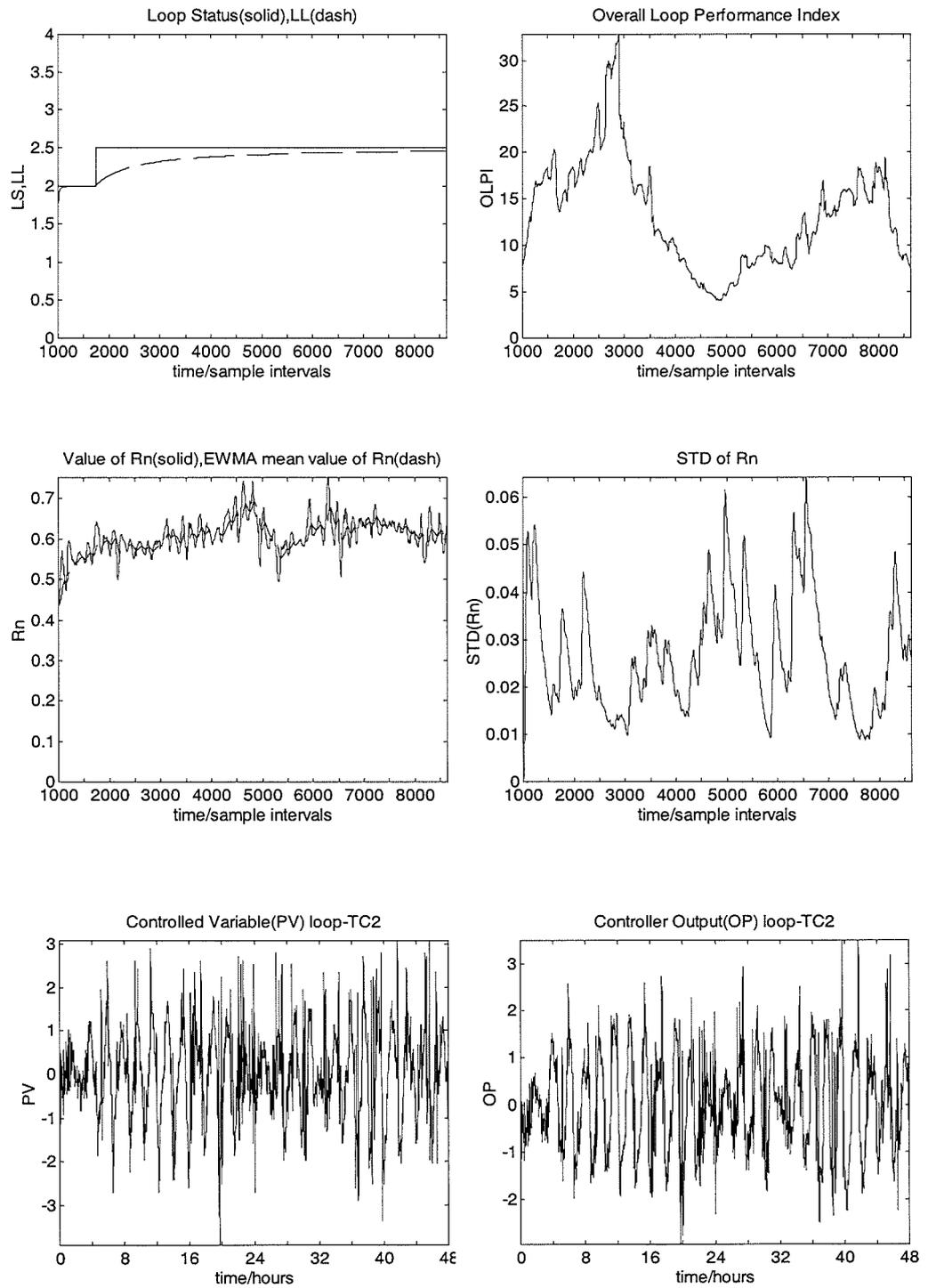


Figure B-8: Results for loop TC2

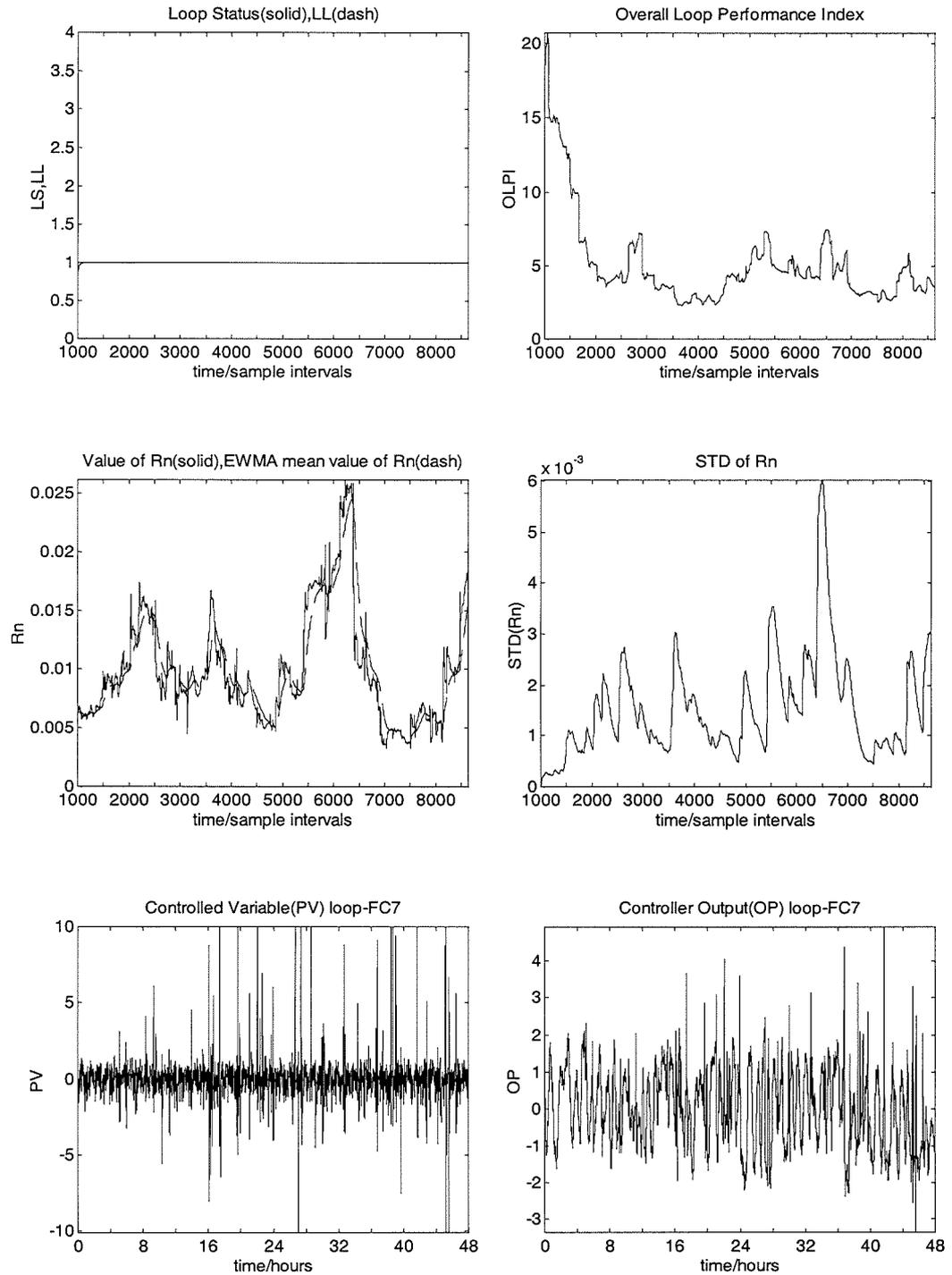


Figure B-9: Results for loop FC7

Appendix C

The Model of 2-Cascaded CSTRs

The simulated model of 2-Cascaded CSTRs used in chapter 5 contains an up stream tank carrying out an endothermic chemical reaction, while down-stream CSTR, an exothermic one. Each CSTR has temperature and level control. The linkage of heating/cooling water recycled between them exhibits the coupled behaviour.

C.1 Up CSTR Tank Model

An irreversible, endothermic reaction is carried out in the up CSTR tank as shown in Figure C-1.

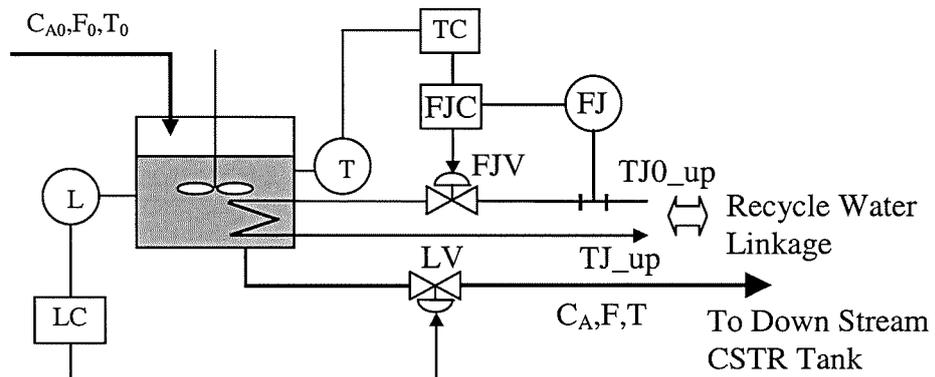


Figure C-1: Simulated up stream CSTR process

The reactant A is fed into the reactor with the initial concentration C_{A0} , flow rate F_0 and temperature T_0 . After the reaction, A leaves the reactor with the final concentration C_A and temperature T , which is viewed as the same as that in the reactor. The reactor temperature, T , is maintained by varying the heating water flow rate, FJ , with the initial temperature TJ_0 , through a heat exchanging jacket installed around the CSTR and, in addition, concentration C_A , outlet heating water temperature TJ are measured. There are three control systems, two of which have single loops, the other has a cascade arrangement. Thus three PI controllers, the level controller LC, the temperature controller TC and the cooling water flow rate controller FJC, are employed to manipulate two valves LV and FJV. Measurement noises and time delays are introduced in control loops.

The dynamic model of the up CSTR is as the following:

$$\frac{dV}{dt} = F_0 - F \quad (C-1)$$

$$\frac{dC_A}{dt} = \frac{F_0}{V}(C_{A0} - C_A) - k_0 e^{\frac{-E}{RT}} C_A \quad (C-2)$$

$$\frac{dT}{dt} = \frac{F_0}{V}(T_0 - T) - \frac{\lambda k_0 e^{\frac{-E}{RT}} C_A}{\rho C_p} - \frac{UA}{V \rho C_p}(T - TJ) \quad (C-3)$$

$$\frac{dTJ}{dt} = \frac{FJ}{V_j}(TJ_0 - TJ) + \frac{UA}{V_j \rho_j C_{pj}}(T - TJ) \quad (C-4)$$

The steady state values & parameter values are listed in Table C-1.

$F = 40 \text{ ft}^3/\text{h}$	$V = 48 \text{ ft}^3$
$C_{A0} = 0.6793 \text{ mol}/\text{ft}^3$	$C_A = 0.5 \text{ mol}/\text{ft}^3$
$T = 530 \text{ }^\circ\text{R}$	$T_{J_up} = 534.93 \text{ }^\circ\text{R}$
$FJ = 49.86 \text{ ft}^3/\text{h}$	$T_0 = 550 \text{ }^\circ\text{R}$
$VJ = 3.85 \text{ ft}^3$	$K_0 = 6.75 \times 10^{11} \text{ h}^{-1}$
$E = 30000 \text{ Btu}/\text{mol}$	$R = 1.99 \text{ Btu}/(\text{mol} \cdot ^\circ\text{R})$
$U_0 = 150 \text{ Btu}/(\text{h} \cdot \text{ft}^2 \cdot ^\circ\text{R})$	$A = 250 \text{ ft}^2$
$TJ0_up = 594.6 \text{ }^\circ\text{R}$	$\lambda = 30000 \text{ Btu}/\text{mol}$
$C_p = 0.75 \text{ Btu}/(\text{lb}_m \cdot ^\circ\text{R})$	$C_{pj} = 1.0 \text{ Btu}/(\text{lb}_m \cdot ^\circ\text{R})$
$\rho = 50 \text{ lb} \cdot \text{mol}/\text{ft}^3$	$\rho_j = 62.3 \text{ lb} \cdot \text{mol}/\text{ft}^3$

Table C-1: Steady state values & parameter values for the up CSTR tank

The parameters for the three controllers are listed in Table C-2.

Controller	Kp	Ti (sec)
TC	20	500
LC	-8.5	60
FJC	3.5	6

Table C-2: Parameters for up CSTR controllers

Time delays used for the three processes are as shown in Table C-3.

Process in loop of	Time Delay (sample intervals)
TC	20
LC	3
FJC	2

Table C-3: Time delays for up CSTR control processes

C.2 Down CSTR Tank Model

The Down stream CSTR tank model is similar to that of up CSTR except that:

- it is an exothermic reaction;
- it has one more outlet flow rate control (FC) as shown in Figure C-2;
- some operating points and parameters are different as listed in Table C-4.

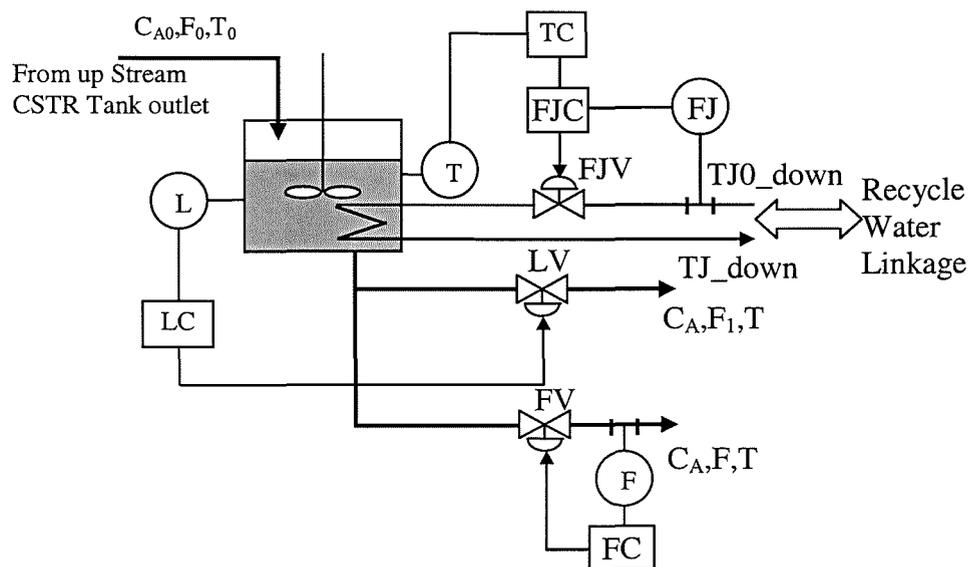


Figure C-2: Simulated down stream CSTR process

$F = 40 \text{ ft}^3/\text{h}$	$V = 48 \text{ ft}^3$
$C_{A0} = 0.50 \text{ mol}/\text{ft}^3$	$C_A = 0.245 \text{ mol}/\text{ft}^3$
$T = 600 \text{ }^\circ\text{R}$	$T_{J_down} = 594.6 \text{ }^\circ\text{R}$
$FJ = 49.86 \text{ ft}^3/\text{h}$	$T_0 = 530 \text{ }^\circ\text{R}$
$VJ = 3.85 \text{ ft}^3$	$K_0 = 7.08 \times 10^{10} \text{ h}^{-1}$
$E = 30000 \text{ Btu}/\text{mol}$	$R = 1.99 \text{ Btu}/(\text{mol} \cdot ^\circ\text{R})$
$U_0 = 150 \text{ Btu}/(\text{h} \cdot \text{ft}^2 \cdot ^\circ\text{R})$	$A = 250 \text{ ft}^2$
$T_{J0_down} = 534.93 \text{ }^\circ\text{R}$	$\lambda = -30000 \text{ Btu}/\text{mol}$
$C_p = 0.75 \text{ Btu}/(\text{lb}_m \cdot ^\circ\text{R})$	$C_{pj} = 1.0 \text{ Btu}/(\text{lb}_m \cdot ^\circ\text{R})$
$\rho = 50 \text{ lb} \cdot \text{mol}/\text{ft}^3$	$\rho_j = 62.3 \text{ lb} \cdot \text{mol}/\text{ft}^3$

Table C-4: Steady state values & parameter values for the down CSTR tank

The parameters for the four controllers are listed in Table C-5.

Controller	Kp	Ti (sec)
TC	10	500
LC	-6	60
FJC	3.5	6
FC	0.35	6

Table C-5: Parameters for down CSTR controllers

Time delays used for the first three processes in the down stream CSTR are the same as those for up CSTR (see Table C-3). The time delay for the process of FC is 2 sample intervals.

C.3 Operating Data for Two CSTRs

Figure C-3 to Figure C-6 show the normal operating data of the up and down stream CSTRs when they are commissioned separately without recycle water linkage affect. Figure C-7 to Figure C-10 illustrate the “*oscillatory plant*” situation when two CSTRs are commissioned simultaneously with the tight linkage of recycle heating/cooling water.

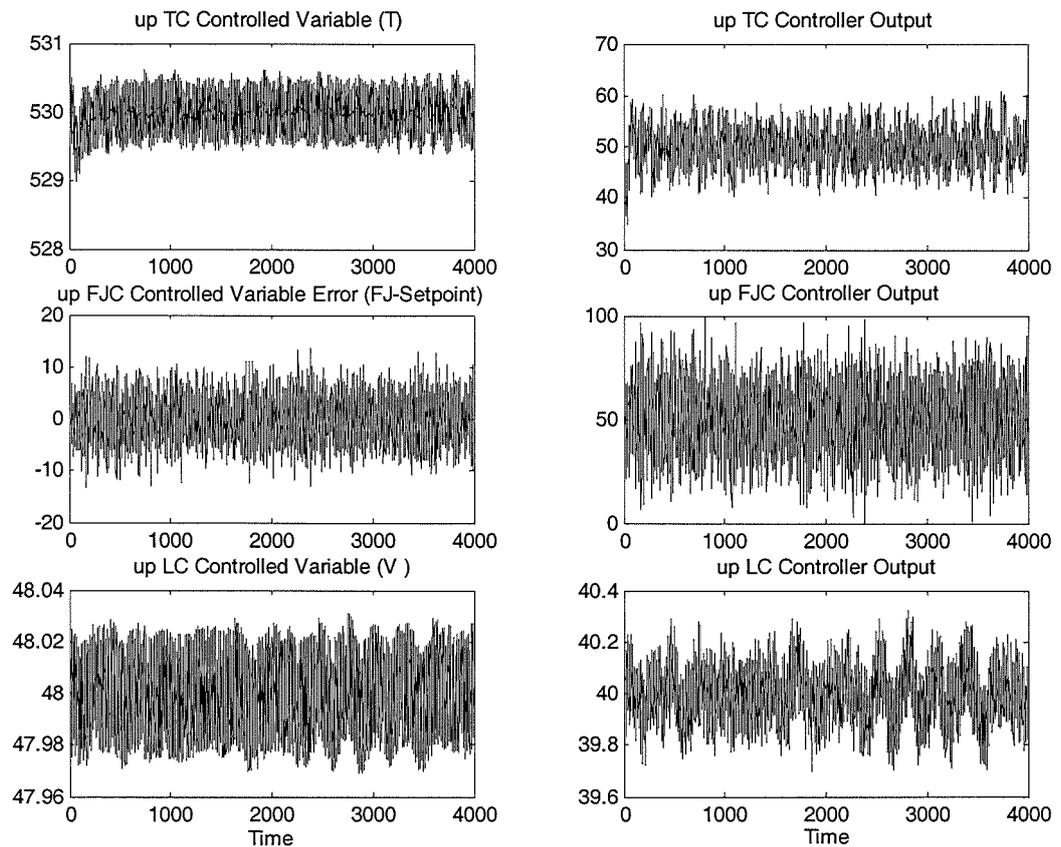


Figure C-3: Controlled variables & controller outputs (up CSTR)

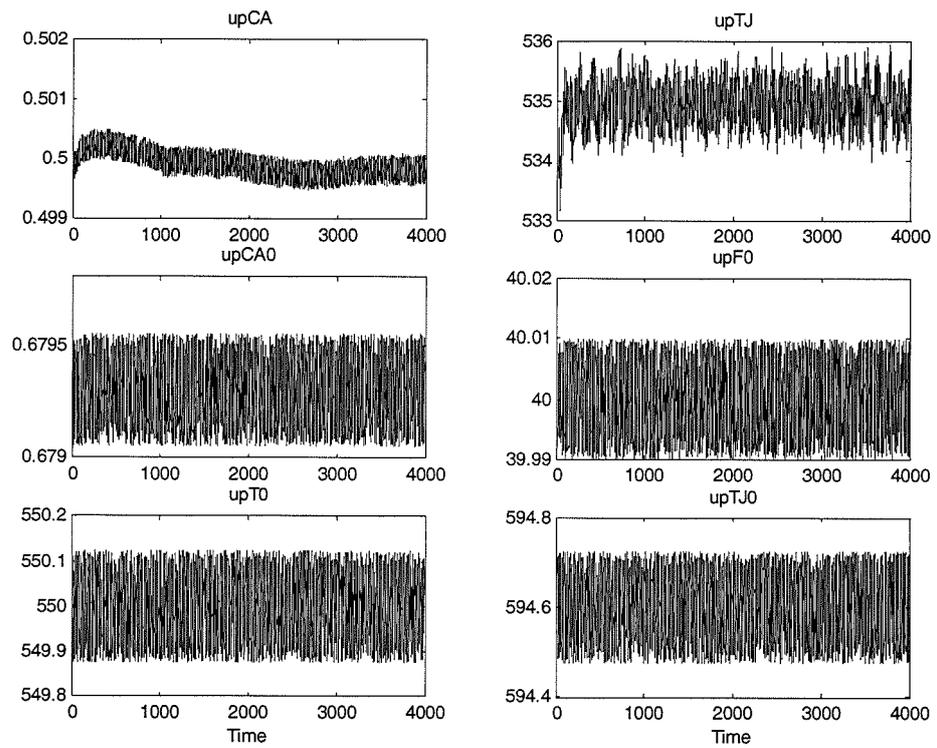


Figure C-4: Other measurements (up CSTR)

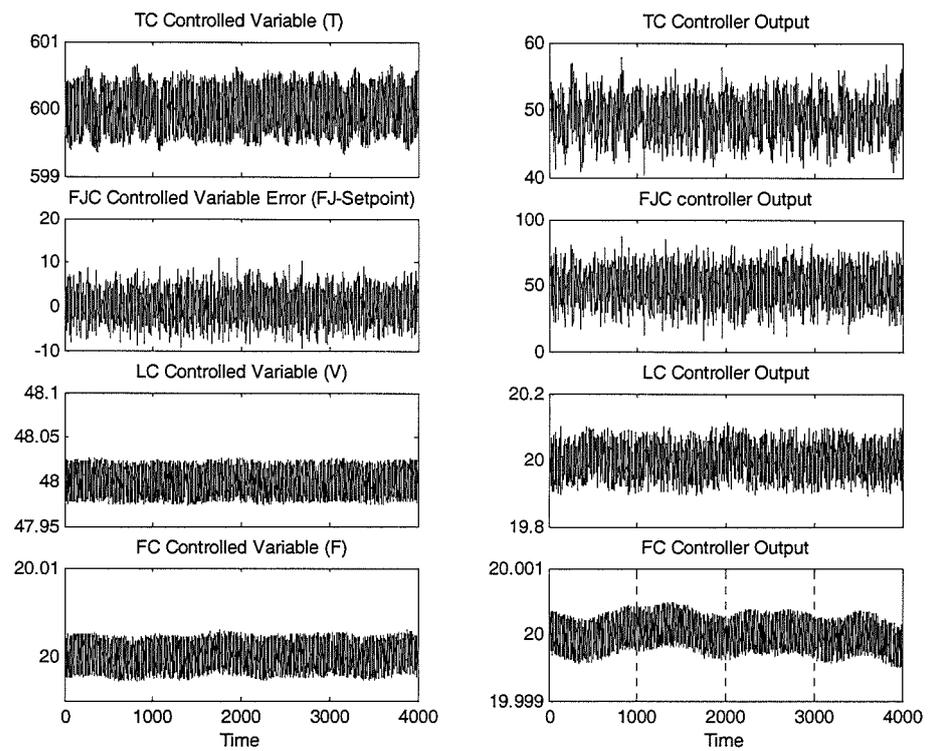


Figure C-5: Controlled variables & controller outputs (down CSTR)

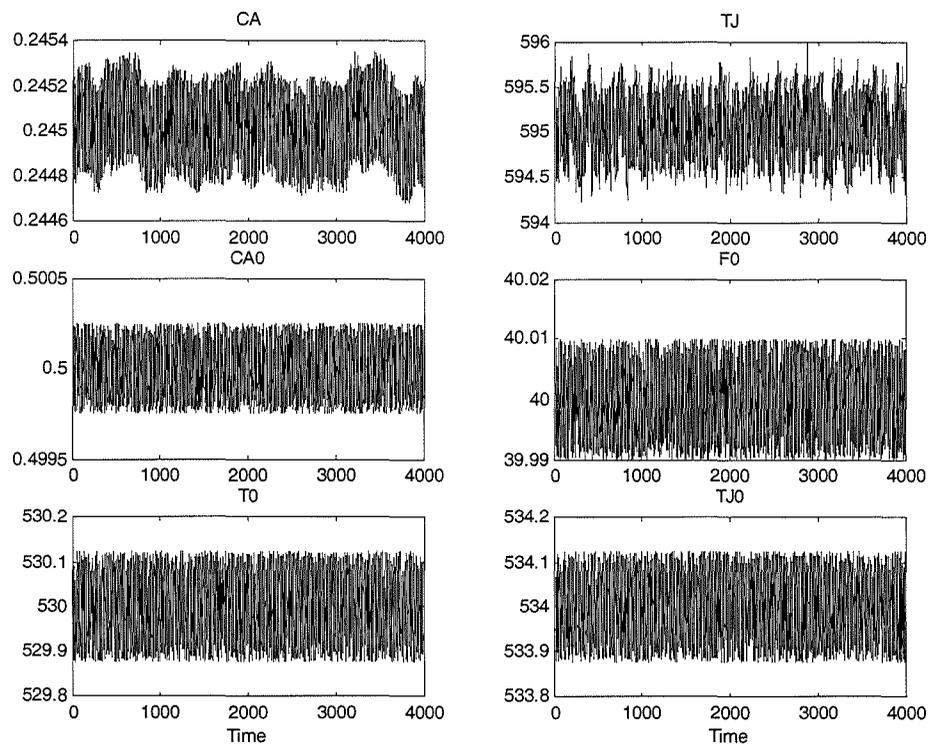


Figure C-6: Other measurements (down CSTR)

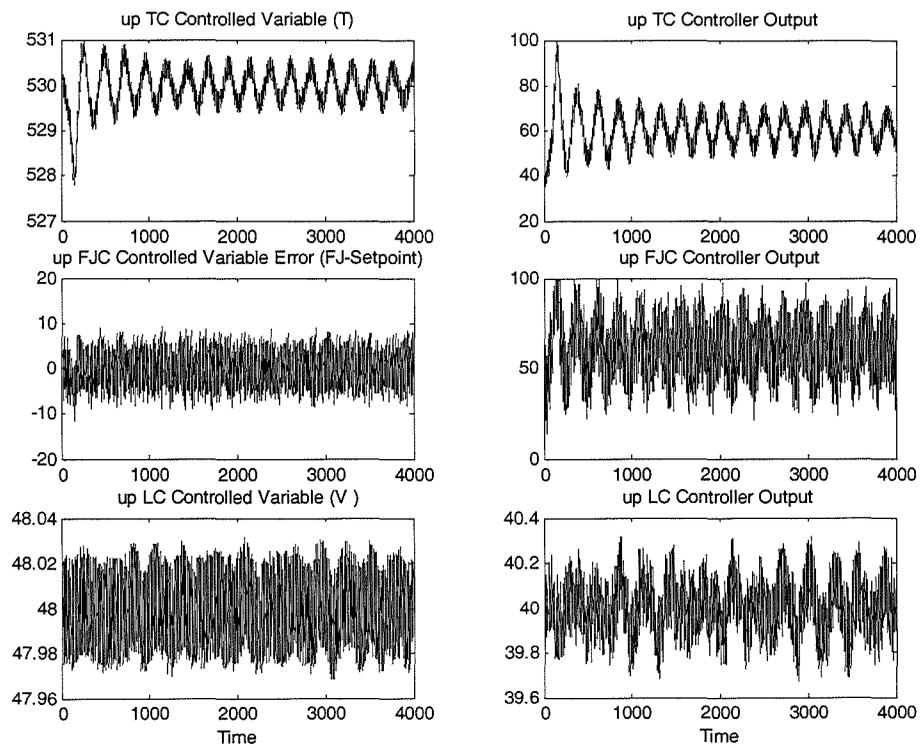


Figure C-7: Controlled variables & controller outputs (up CSTR, linkage)

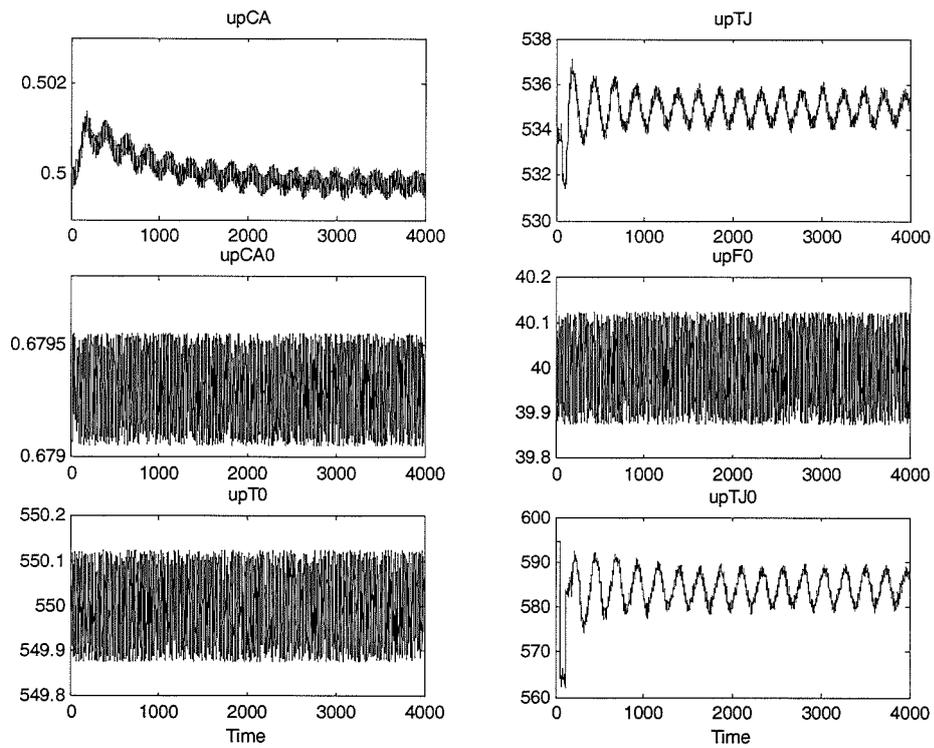


Figure C-8: Other measurements (up CSTR, linkage)

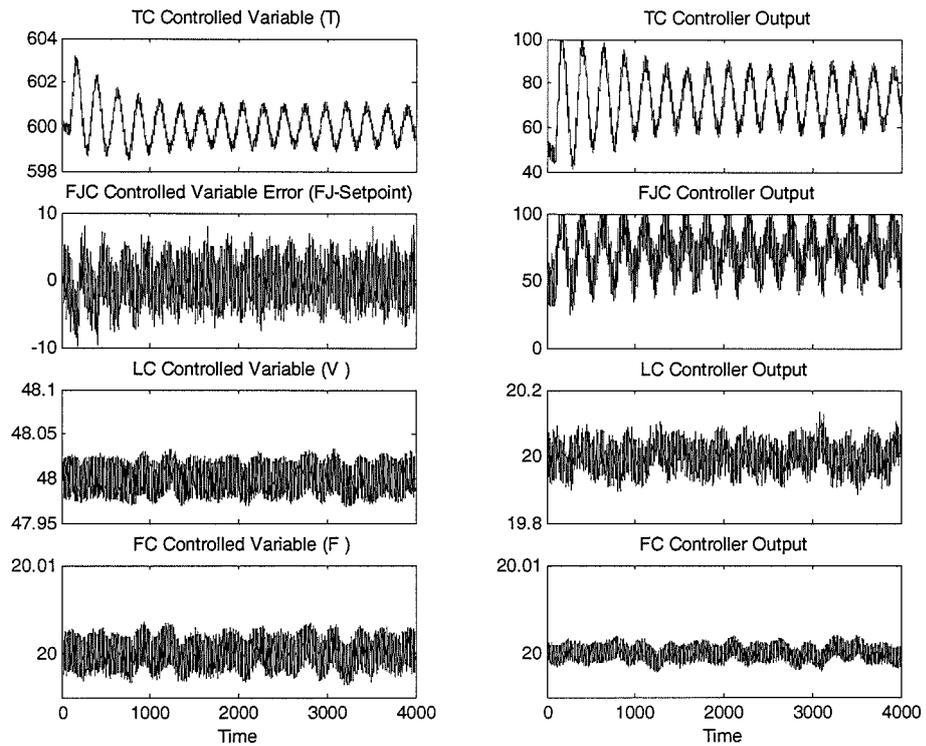


Figure C-9: Controlled variables & controller outputs (down CSTR, linkage)

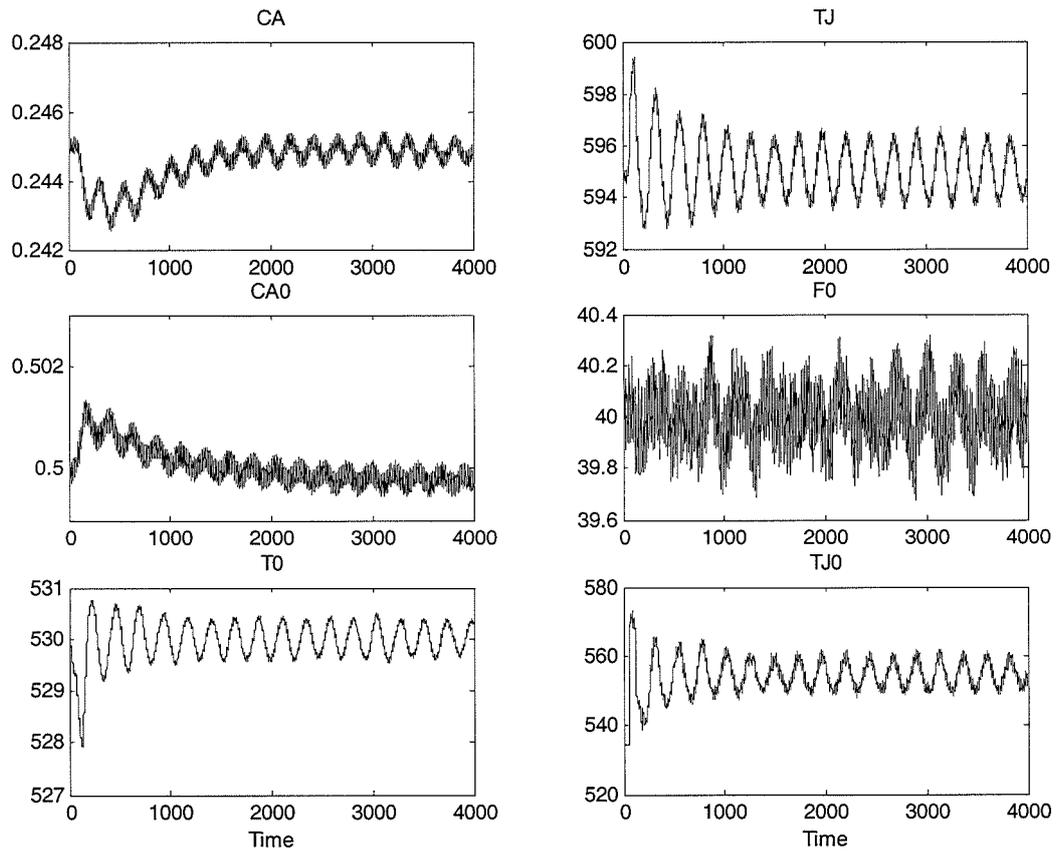


Figure C-10: Other measurements (down CSTR, linkage)

Appendix D

The Kurtosis of Some Typical Power Spectra

The definition of the normalised kurtosis of a random variable x is:

$$kurt(x) = \frac{\mu_4 - 3\mu_2^2}{\mu_2^2} \quad (\text{D.1})$$

where μ_4 is the fourth-order central moment of x , and μ_2 is the 2nd-order central moment, i.e., variance of x .

Lemma 1:

The kurtosis of the power spectrum of a finite sample of Gaussian noise is 6.

Proof 1:

Let $[x_1, x_2 \dots x_{\tilde{N}}]$ be the zero mean Gaussian noise series and its \tilde{N} -point FFT be the complex vector C where $C = R + jI$. Its elements are defined by:

$$C(k) = \sum_{m=1}^{\tilde{N}} x_m e^{-j\left(\frac{2\pi(m-1)(k-1)}{\tilde{N}}\right)}, k = 1 \dots \tilde{N} \quad (\text{D.2})$$

Then C has a complex Normal distribution, i.e., both the real part R and the imaginary part I will have Normal distributions with zero means and the same variances, say σ^2 , and the real and imaginary parts are independent (Ljung 1999).

The 2-sided power spectrum of Gaussian noise is then given by:

$$P_{xx}(k) = \frac{C(k)^* C(k)}{\tilde{N}} = \frac{R(k)^2 + I(k)^2}{\tilde{N}}, k = 1 \dots \tilde{N}, \quad (\text{D.3})$$

which has a special Gamma distribution that is exponential:

$$f(t) = \frac{1}{\beta^\alpha \Gamma(\alpha)} t^{\alpha-1} e^{-\frac{t}{\beta}} = \frac{1}{\beta} e^{-\frac{t}{\beta}} \quad (\text{D.4})$$

with $\alpha = 1$ and $\beta = \frac{2\tilde{N}}{\sigma^2}$, where $\Gamma(\cdot)$ is the Gamma function (Fisz 1963).

The moment generating function for the exponential distribution is:

$$M(t) = (1 - \beta t)^{-1} \quad (\text{D.5})$$

and the r^{th} moment can be given by $\mu'_r = \left. \frac{d^r M(t)}{dt^r} \right|_{t=0}$, $r = 1, 2, 3, 4 \dots$, e.g.

$$\mu'_1 = \beta, \mu'_2 = 2\beta^2, \mu'_3 = 6\beta^3, \mu'_4 = 24\beta^4.$$

It is known that any central-moment μ_r can be evaluated from moments of order up to r . for example, $\mu_2 = \mu'_2 - (\mu'_1)^2$ and $\mu_4 = \mu'_4 - 4\mu'_1\mu'_3 + 6(\mu'_1)^2\mu'_2 - 3(\mu'_1)^4$ (Papoulis 1991).

The kurtosis of P_{xx} is then

$$\text{kurt}(P_{xx}) = \frac{\mu_4 - 3\mu_2^2}{\mu_2^2} = \frac{9\beta^4 - 3(\beta^2)^2}{(\beta^2)^2} = 6 \quad (\text{D.6})$$

Lemma 2:

The kurtosis of the power spectrum of a pure sine wave, $x = A \sin(\omega_x t)$, is $\frac{\tilde{N}}{2} - 3$,

where \tilde{N} is the number of FFT data points.

Proof 2:

The 2-sided power spectrum of x is given by:

$$P_{xx} = U\delta(\omega - \omega_x) + U\delta(\omega - (\omega_s - \omega_x)) \quad (\text{D.7})$$

where $\delta(\cdot)$ is the *delta* function, ω_s is sampling frequency, and $U = \frac{\tilde{N}}{4}A^2$ is the magnitude of the power spectrum peak. The second- and fourth-order sample central moments can be calculated according to the natural moment estimators:

$$\mu_2 = \frac{2(U - \frac{2U}{\tilde{N}})^2}{\tilde{N}} \quad (\text{D.8})$$

$$\mu_4 = \frac{2(U - \frac{2U}{\tilde{N}})^4}{\tilde{N}} \quad (\text{D.9})$$

The kurtosis is then:

$$kurt(P_{xx}) = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\frac{2(U - \frac{2U}{\tilde{N}})^4}{\tilde{N}}}{\left(\frac{2(U - \frac{2U}{\tilde{N}})^2}{\tilde{N}}\right)^2} - 3 = \frac{\tilde{N}}{2} - 3 \quad (\text{D.10})$$

Note that the kurtosis is independent of the signal frequency ω_x , and has a linear relationship with the number of FFT data points.

Lemma 3:

The kurtosis of the power spectrum of a deterministic trend x with several significant harmonics: $x = \sum_{i=1}^k A_i \sin(\omega_i t)$, where $\omega_i \neq \omega_j (\forall i \neq j)$, has a lower bound of $\frac{\tilde{N}}{2k} - 3$, and an upper bound of $\frac{\tilde{N}}{2} - 3$, where \tilde{N} is the number of FFT data points.

Proof 3:

The 2-side power spectrum of x is given by:

$$P_{xx} = \sum_{i=1}^k [\alpha_i \delta(\omega - \omega_i) + \alpha_i \delta(\omega - (\omega_s - \omega_i))] \quad (\text{D.11})$$

where $\delta(\cdot)$ is the *delta* function, ω_s is sampling frequency, and $\alpha_i (i=1 \dots k)$ is the magnitude of the relevant power spectrum peak. The second- and fourth-order sample central moments are:

$$\mu_2 = \frac{\sum_{i=1}^k 2 \left(\alpha_i - \frac{\sum_{i=1}^k 2\alpha_i}{\tilde{N}} \right)^2}{\tilde{N}} = \frac{2 \sum_{i=1}^k B_i}{\tilde{N}} \quad (\text{D.12})$$

$$\mu_4 = \frac{\sum_{i=1}^k 2 \left(\alpha_i - \frac{\sum_{i=1}^k 2\alpha_i}{\tilde{N}} \right)^4}{\tilde{N}} = \frac{2 \sum_{i=1}^k B_i^2}{\tilde{N}} \quad (\text{D.13})$$

where $B_i = \left(\alpha_i - \frac{\sum_{i=1}^k 2\alpha_i}{\tilde{N}} \right)^2$, $B_i > 0$, $(i=1 \dots k)$

The kurtosis is then:

$$kurt(P_{xx}) = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\frac{2 \sum_{i=1}^k B_i^2}{\tilde{N}}}{\left(\frac{2 \sum_{i=1}^k B_i}{\tilde{N}} \right)^2} - 3 = \frac{\tilde{N}}{2} \left(\frac{\sum_{i=1}^k B_i^2}{\left(\sum_{i=1}^k B_i \right)^2} \right) - 3 \quad (\text{D.14})$$

It is clear that: $kurt(P_{xx}) \leq \frac{\tilde{N}}{2} - 3$, and the equality will hold only when $k = 1$.

The low bound, i.e., $Kurto(P_{xx}) \geq \frac{\tilde{N}}{2k} - 3$ can be found by considering the inequality equation (D.15):

$$\left(\frac{\sum_{i=1}^k B_i}{k} \right)^2 \leq \frac{\sum_{i=1}^k B_i^2}{k}, \quad (\text{D.15})$$

For the low bound, the equality will only hold when $B_1 = B_2 = \dots B_k$, i.e., all dominant sine wave components have the same height of their peaks in the power spectrum, i.e., the same energy.

Demonstration

Figure D-1 demonstrates the kurtosis of the power spectrum of a deterministic trend containing 2 pure sine wave components by varying the B_1 percentage value of a sine wave component, i.e., $\frac{B_1}{B_1 + B_2}$. The frequencies of the two sinusoids are 1/40 and 1/100 Hz. The power spectrum was calculated by FFT and the number of FFT

data points was $\tilde{N} = 1000$. The theoretical lower and upper bounds were calculated as [247, 497]. It can be seen that the estimated result fits the theoretical one very well, and the low bound is exactly reached when the two components have the same percentage, i.e., $B_1 = B_2 = 0.5$, which implies the equal energy of these two sine wave components.

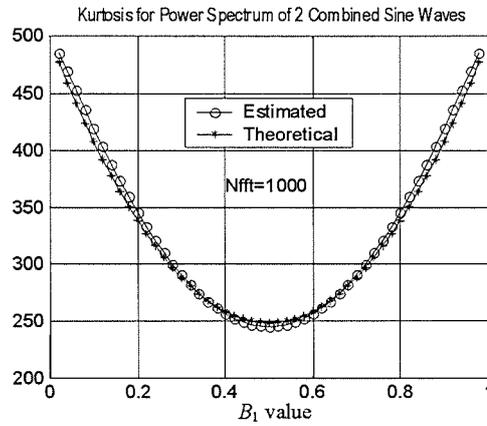


Figure D-1: Kurtosis of the power spectrum of a deterministic trend with 2 sine waves

