Multivariate Statistical Process Monitoring Using Classical Multidimensional Scaling

Prepared by:

Mohd Yusri Mohd Yunus

A Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

School of Chemical Engineering and Advanced Materials

Newcastle University

UK

May 2012

ABSTRACT

A new Multivariate Statistical Process Monitoring (MSPM) system, which comprises of three main frameworks, is proposed where the system utilizes Classical Multidimensional Scaling (CMDS) as the main multivariate data compression technique instead of using the linear-based Principal Component Analysis (PCA). The conventional method which usually applies variance-covariance or correlation measure in developing the multivariate scores is found to be inappropriately used especially in modelling nonlinear processes, where a high number of principal components will be typically required. Alternatively, the proposed method utilizes the inter-dissimilarity scales in describing the relationships among the monitored variables instead of variance-covariance measure for the multivariate scores development. However, the scores are plotted in terms of variable structure, thus providing different formulation of statistics for monitoring. Nonetheless, the proposed statistics still correspond to the conceptual objective of Hotelling's T^2 and Squared Prediction Errors (SPE). The first framework corresponds to the original CMDS framework, whereas the second utilizes Procrustes Analysis (PA) functions which is analogous to the concept of loading factors in PCA for score projection. Lastly, the final framework employs dynamic mechanism of PA functions as an alternative for enhancing the procedures of the second approach. A simulated system of Continuous Stirred Tank Reactor with Recycle (CSTRwR) has been chosen for the demonstration and the fault detection results were comparatively analyzed to the outcomes of PCA on the grounds of false alarm rates, total number of detected cases and also total number of fastest detection cases. The last two performance factors are obtained through fault detection time. The overall outcomes show that the three CMDS-based systems give almost comparable performances to the linear PCA based monitoring systemwhen dealing the abrupt fault events, whereas the new systems have demonstrated significant improvement over the conventional method in detecting incipient fault cases. More importantly, this monitoring accomplishment can be efficiently executed based on lower compressed dimensional space compared to the PCA technique, thus providing much simpler solution. All of these evidences verified that the proposed approaches are successfully developed conceptually as well as practically for monitoring while complying fundamentally with the principles and technical steps of the conventional MSPM system.

ACKNOWLEDGEMENT

First of all, I would like to thank to both of my respectful supervisors Dr. Jie Zhang and Professor Julian Morris who have relentlessly provided me the guidance, constructive ideas and invaluable advice, enabling me to achieve the objective of this thesis. I am also very grateful to my parents, my wife, my daughter and all of my family members for their endless inspiration, kindness, understanding as well as patience. I am also abundantly indebted to many people, particularly in SAGE as well as SCEAM of Newcastle University, UK, and also Universiti Malaysia Pahang (UMP) as well as Ministry of Higher Educational Malaysia (MoHE), whom have helped me in various ways, to start, continue and complete the study. May God bless us all.

" The goal of the journey is the journey itself...poetically put, it is a journey that takes us far away, and back to ourselves...Humility is my table, respect is my garment, empathy is my food and curiosity is my drink. As for love, it has a thousand names and is by my side at every window."

(Tariq Ramadan, The Quest For Meaning, 2010).

TABLE OF CONTENTS

CHAPTERS	TITLES		PAGES
	TOPIC PAGE		i
	ABSTRACT		ii
	ACKNOWLEDGEMENT		iii
	TABLE OF CONTENTS		iv
	LIST OF TABLES		viii
	LIST OF FIGURES		x
	LIST OF ABBREVIATIONS		xvi
	LIST OF SYMBOLS		xviii
CHAPTER 1	**INTRODUCTION**		1
	1.1	Background	1
	1.2	Motivation	4
	1.3	Aim and Objectives	8
	1.4	Scopes	9
	1.5	Contributions	10
	1.6	Thesis Organization	12
CHAPTER 2	**MULTIVARIATE STATISTICAL PROCESS MONITORING**		14
	2.1	Introduction	14
	2.2	Fundamentals of Conventional MSPM System	15
		2.2.1 Phase I: Off-line Modelling and Monitoring	17

		2.2.2 Phase II: On-line Monitoring	23
	2.3.	MSPM Issues and Extensions	27
		2.3.1 Issue I: Monitoring Statistics	28
		2.3.2 Issue 2: Monitoring Limits	28
		2.3.3 Issue 3: Multivariate techniques	30
		2.3.4 Issue 4: Fault Diagnosis	36
		2.3.5 Issue 5: Data Formulation	38
		2.3.6 Issue 6: Automation System	39
	2.4.	MDS as An Alternative Solution for MSPM System	39
		2.4.1 Conceptual Background of MDS	40
		2.4.2 Mathematical Fundamentals of MDS	44
		2.4.3 Connections between PCA and MDS	50
		2.4.4 Previous Works on MDS-based Monitoring Systems	52
		2.4.5 Justification of Applying MDS in The MSPM Framework	54
	2.5	Summary	55
CHAPTER 3	**CASE STUDIES AND PCA-BASED MONITORING PERFORMANCES**		56
	3.1	Introduction	56
	3.2	CSTRwR System	56
		3.2.1 Process Descriptions	56
		3.2.2 NOC Samples and PCA Monitoring Performances	57
		3.2.3 Fault Cases PCA Monitoring Performances	61
	3.3	Summary	71
CHAPTER 4	**FRAMEWORK I: MDS-BASED MSPM SYSTEM USING MOVING WINDOW CMDS PROJECTION**		72
	4.1	Introduction	72
	4.2	Methodology	72
		4.2.1 Phase I Procedures	73
		4.2.2 Phase II Procedures	77

	4.3	Results and Analysis	79
	4.4	Results Discussion	100
		4.4.1 The Impact of Dissimilarity Measures on The Monitoring Outcomes	100
		4.4.2 The Impact of using New Monitoring Statistics on The Monitoring Outcomes	100
		4.4.3 The Impact of Applying Various Window Settings on The Monitoring Outcomes	101
		4.4.4 The Impact of Applying Smaller Dimensionalities on The Monitoring Outcomes	102
	4.5	Summary	103
CHAPTER 5		**FRAMEWORK II: MDS-BASED MSPM SYSTEM USING MOVING WINDOW CMDS PROCRUSTES ANALYSIS PROJECTION**	104
	5.1	Introduction	104
	5.2	Methodology	105
		5.2.1 Phase I Procedures	105
		5.2.2 Phase II Procedures	109
	5.3	Results and Analysis	111
	5.4	Results Discussion	132
		5.4.1 The Impact of Dissimilarity Measures on The Monitoring Outcomes	132
		5.4.2 The Impact of using New Monitoring Statistics on The Monitoring Outcomes	132
		5.4.3 The Impact of Applying Various Window Settings on The Monitoring Outcomes	133
		5.4.4 The Impact of Applying Smaller Dimensionalities on The Monitoring Outcomes	133

	5.5	Summary	134
CHAPTER 6	**FRAMEWORK III: MDS-BASED MSPM SYSTEM USING MOVING WINDOW CMDS DYNAMICAL PROCRUSTES ANALYSIS PROJECTION**		135
	6.1	Introduction	135
	6.2	Methodology	136
		6.2.1 Phase I Procedures	137
		6.2.2 Phase II Procedures	138
	6.3	Results and Analysis	141
	6.4	Results Discussion	154
		6.4.1 The Impact of Dissimilarity Measures on The Monitoring Outcomes	154
		6.4.2 The Impact of using New Monitoring Statistics on The Monitoring Outcomes	154
		6.4.3 The Impact of Applying Various Window Settings on The Monitoring Outcomes	155
		6.4.4 The Impact of Applying Smaller Dimensionalities on The Monitoring Outcomes	156
	6.5	Summary	157
CHAPTER 7	**CONCLUSIONS AND RECOMMENDATIONS**		158
	7.1	Conclusions	158
	7.2	Recommendations for further works	162
	REFERENCES		164
	APPENDIX A		172
	APPENDIX B		173
	APPENDIX C		176

LIST OF TABLES

TABLE NO.	TITLES	PAGES
Table 2.1	Examples of industrial applications of MSPM	27
Table 2.2	Categories of MSPM Issues and Extensions Reviewed in This Study	27
Table 2.3	Extensions of PCA-based monitoring techniques	31
Table 2.4	Dissimilarity matrix based on Euclidean scales used in Figure 2.2	48
Table 3.1	List of variables in the CSTRwR system for monitoring	57
Table 3.2	List of abnormal operations in CSTRwR	61
Table 4.1	Fault detection times of monitoring systems based on sCMDS and PCA for abrupt fault cases	82
Table 4.2	Fault detection times of monitoring systems based on sCMDS and PCA for incipient fault cases	83
Table 5.1	Fault detections times of monitoring systems based on CMDS-PA and PCA for abrupt fault cases	113
Table 5.2	Fault detection times of monitoring systems based on CMDS-PA and PCA for incipient fault cases	114
Table 5.3	Monitoring limits specified for S_r based on sCMDS and CMDS-PA methods	132
Table 6.1	Results of FDT of abrupt fault cases based on CMDS-dPA for CSTRwR	142
Table 6.2	Results of FDT of incipient fault cases based on CMDS-dPA for CSTRwR	143

Table 6.3 Monitoring limits specified for S_{m2} based on Euclidean scale and window size 5 154

Table 6.3 Monitoring limits specified for S_{r} based on Euclidean scale and window size 5 155

LIST OF FIGURES

FIGURE NO.	TITLES	PAGES
Figure 2.1	Conventional MSPM framework	17
Figure 2.2	A numerical example for CMDS	47
Figure 2.3	The reproduction of CMDS scores that demonstrated in Figure 2.2	48
Figure 3.1	CSTRwR system	57
Figure 3.2	Accumulated data variations explained by different PCs for the CSTRwR	58
Figure 3.3	Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) on the original NOC data	59
Figure 3.4	Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) on the testing NOC data	60
Figure 3.5	The behaviours of fault number 6 based on trends of variables 1 (top charts), 7 (middle charts) and 13 (bottom charts)	62
Figure 3.6	Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F6a data of CSTRwR	63
Figure 3.7	Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F6i data of CSTRwR	64
Figure 3.8	The behaviours of fault number 9 based on trends of variables 1 (top charts), 13 (middle charts) and 12 (bottom charts) of CSTRwR system for abrupt (left diagrams) and incipient (right diagrams) fault categories	65
Figure 3.9	Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F9a data of CSTRwR	66

Figure 3.10 Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F9i data of CSTRwR 67

Figure 3.11 The behaviours of fault number 11 based on trends of variables 9 (top charts), and 8 (bottom charts) of CSTRwR system for abrupt (left diagrams) and incipient (right diagrams) fault categories 68

Figure 3.12 Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F11a data of CSTRwR 69

Figure 3.13 Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F11i data of CSTRwR 70

Figure 4.1 CMDS-based MSPM framework 73

Figure 4.2 Illustration of S_{m1} based on the plots of NOC scores vs MWOS-NOC scores (left diagram); NOC scores vs MWOS-fault scores (right diagram) 74

Figure 4.3 Illustration of S_r based on the plots of NOC scores vs MWOS-NOC scores (left diagram); NOC scores vs MWOS-fault scores (right diagram) 76

Figure 4.4 Accumulated portion of data variation explained by the dimensions of the CMDS model of $\mathbf{X_{NOC1}}$ 79

Figure 4.5 Monitoring progression of S_{m1}(left) and S_r (right) on F6a based on sCMDS models using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 85

Figure 4.6 Monitoring progression of S_{m1}(left) and S_r (right) on F6i based on sCMDS models using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 86

Figure 4.7 Contribution plots of S_{m1} (top) and S_r (bottom) for F6a with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C) 88

Figure 4.8 Contribution plots of S_{m1} (top) and S_r (bottom) for F6i with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C) 89

Figure 4.9 Monitoring progression of S_{m1}(left) and S_r (right) on F9a based on sCMDS using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 91

Figure 4.10 Monitoring progression of S_{m1}(left) and S_r (right) on F9i based on sCMDS using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 92

Figure 4.11 Contribution plots of S_{m1} (top) and S_r (bottom) for F9a with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C) 93

Figure 4.12 Contribution plots of S_{m1} (top) and S_r (bottom) for F9i with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C) 94

Figure 4.13 Monitoring progression of S_{m1}(left) and S_r (right) on F11a based on sCMDS using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 96

Figure 4.14 Monitoring progression of S_{m1}(left) and S_r (right) on F11i based on sCMDS using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 97

Figure 4.15 Contribution plots of S_{m1} (top) and S_r (bottom) for F11a with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C) 98

Figure 4.16 Contribution plots of S_{m1} (top) and S_r (bottom) for F11i with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C) 99

Figure 4.17 NOC scores (‘o’) vs MWOS-new scores (‘x’) 101

Figure 5.1 CMDS-PA-based MSPM framework 105

Figure 5.2 Illustration of S_{m2} based on the plots of NOC scores vs MWOS-NOC scores (left diagram); NOC scores vs MWOS-fault scores (right diagram) 108

Figure 5.3 Monitoring progression of S_{m2}(left) and S_r (right) for F6a from monitoring systems based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 117

Figure 5.4 Monitoring progression of S_{m2} (left) and S_r (right) for F6i from monitoring systems based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 118

Figure 5.5 Conventional contribution plots (left) and differential contribution plots (right) for F6a from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C) 119

Figure 5.6 Conventional contribution plots (left) and differential contribution plots (right) for F6i from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C) 120

Figure 5.7 Monitoring progression of S_{m2} (left) and S_r (right) for F9a from monitoring systems based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 123

Figure 5.8 Monitoring progression of S_{m2} (left) and S_r (right) for F9i from monitoring systems based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 124

Figure 5.9 Conventional contribution plots (left) and differential contribution plots (right) for F9a from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C) 125

Figure 5.10 Conventional contribution plots (left) and differential contribution plots (right) for F9i from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C) 126

Figure 5.11 Monitoring progression of S_{m2}(left) and S_r (right) for F11a from monitoring systems based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 128

Figure 5.12 Monitoring progression of S_{m2}(left) and S_r (right) for F11i from monitoring systems based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom) 129

Figure 5.13 Conventional contribution plots (left) and differential contribution plots (right) for F11a from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C) 130

Figure 5.14 Conventional contribution plots (left) and differential contribution plots (right) for F11i from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C) 131

Figure 6.1 CMDS-dPA-based MSPM framework 136

Figure 6.2 Monitoring progression of S_{m2}(left) and S_r (right) for F6a based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom) 146

Figure 6.3 Monitoring progression of S_{m2}(left) and S_r (right) for F6i based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom) 146

Figure 6.4 Differential contribution plots for F6a from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B) 147

Figure 6.5 Differential contribution plots for F6i from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B) 147

Figure 6.6 Monitoring progression of S_{m2}(left) and S_r (right) for F9a based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom) 149

Figure 6.7 Monitoring progression of S_{m2}(left) and S_r (right) for F9i based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom) 149

Figure 6.8 Differential contribution plots for F9a from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B) 150

Figure 6.9 Differential contribution plots for F9i from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B) 150

Figure 6.10 Monitoring progression of S_{m2}(left) and S_r (right) for F11a based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom) 152

Figure 6.11 Monitoring progression of S_{m2}(left) and S_r (right) for F11i based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom) 152

Figure 6.12 Differential contribution plots for F11a from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B) 153

Figure 6.13 Differential contribution plots for F11i from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B) 153

LIST OF ABBREVIATIONS

APC	-	Automatic process control.
CMDS	-	Classical scaling.
CMDS-dPA	-	CMDS-dynamical Procrustes analysis method.
CMDS-PA	-	CMDS-Procrustes analysis method.
CSTRwR	-	Continuous stirred tank reactor with recycle.
CVA	-	Canonical variate analysis.
DDQI	-	Data driven quality improvement.
DISSIM	-	Dissimilarity index.
FAR	-	False alarm analysis.
FCA	-	Factor analysis.
FDT	-	Fault detection time.
ICA	-	Independent component analysis.
JPMC	-	Jackson's process monitoring criteria.
KBS	-	Knowledge-based system.
KDE	-	Kernel distribution estimation.
KPCA	-	Kernel PCA.
M-CUSU M	-	Multivariate cumulative sum.
MDS	-	Multidimensional scaling.
MDV	-	Multidimensional visualization.
MEWMA	-	Multivariate exponential weight moving average.
MLR	-	Multiple linear regression.
MSPC	-	Multivariate statistical process control.
MSPM	-	Multivariate statistical process monitoring.

MVA	-	Multivariate analysis.
MWOS	-	Moving window observation samples.
MW-PCA	-	Moving window PCA.
ND	-	No detection.
NMDS	-	Non-metric MDS.
NOC	-	Normal operating condition.
PA	-	Procrustes analysis.
PARAFAC	-	Parallel factor analysis.
PC	-	Principal components.
PCA	-	Principal component analysis.
PCarrA	-	Partial correlation analysis.
PDF	-	Probability distribution function.
PLS	-	Partial least square.
SCE	-	Single channel event.
sCMDS	-	Standard CMDS method.
SD	-	Singular decomposition.
SPC	-	Statistical process control.
SPE	-	Squared prediction errors.
SPM	-	Statistical process monitoring.
TE	-	Tennessee Eastman.

LIST OF SYMBOLS

a	-	Number of PCs retained in the PCA model.
a_{ij}	-	Factor loadings in factor analysis algorithms.
$\|\mathbf{a}'\|$	-	Scalar product.
A	-	Transformation matrix of **u** for **y** prediction in the CVA algorithms.
b_{jk}	-	Scalar product between point *'j'* and *'k'*.
b_{ij}	-	Principal component loadings in factor analysis algorithms.
B	-	Transformation matrix of **w** for **y** prediction the in CVA algorithms.
$\mathbf{B}_{\Delta}$	-	Double-centred dissimilarity matrix.
$\mathbf{B}_{\mathbf{E}}$	-	Double-centred dissimilarity matrix of reconstructed scores.
c_{mm}	-	A particular element in variance-covariance matrix at row *'m'* and column *'m'*.
$Cont_{T^2}(x_i)$	-	Contribution of T^2 for variable *'i'*.
C	-	City-block distance.
C	-	Variance-covariance matrix.
$\mathbf{C}_{\mathbf{PA}}$	-	Variance-covariance matrix in the PA algorithms.
d	-	Inter-scores distance.
$d_{dPA(ij)}$	-	Dissimilarity element of CMDS-dPA scores.
$d_{PA(ij)}$	-	Dissimilarity element of CMDS-PA scores.
$d_{sCMDS(ij)}$	-	Dissimilarity element of sCMDS scores.
dc	-	Differential contribution.
D	-	Eigenvalue matrix for SVD.
E	-	Euclidean distance.
E	-	Noise vector for **y** prediction in the CVA algorithms.
$\tilde{\mathbf{e}}_{\mathbf{i}}$	-	The *'i'*th row vector in residual matrix.
E	-	Residual matrix in the ICA algorithms.
$\tilde{\mathbf{E}}$	-	Residual matrix.

$\mathbf{F}$	-	Transformation matrix of **x** for **x** prediction in the CVA algorithms.
$\mathbf{F}_i$	-	Factor analysis scores.
$F_{a,n-a,\alpha}$	-	*F* distributional index with *A* and *n-A* degrees of freedom at α confident limit.
$\mathbf{G}$	-	Transformation matrix of **u** for **x** prediction in the CVA algorithms.
$\mathbf{H}$	-	Transformation matrix of **x** for **y** prediction in the CVA algorithms.
$\mathbf{I}$	-	Identity matrix.
k	-	Sampling time.
l	-	Strain parameter.
m	-	Number of variables.
$\overline{m}$	-	Means of MDS statistics.
n	-	Number of samples.
$\mathbf{q}_i$	-	Normalized eigenvector for dimension *'i'*.
$\mathbf{q}_i^*$	-	Original eigenvector for dimension *'i'*.
$\mathbf{Q}$	-	Eigenvector matrix of minor product moment.
p	-	Number of dimensions selected.
$p_{i,j}$	-	i^{th} element for principal component *j*.
$\mathbf{p_m}$	-	The *'m'*th column vector in PC matrix.
$\mathbf{P}$	-	PC scores matrix.
$\mathbf{P_{PA}}$	-	Eigenvector matrix in the PA algorithms.
$r_{yx\|z1,z2,z3,...,zn}$	-	Partial coefficient of *y* and *x* after when z_1 until z_n are controlled.
$r_{yx\|z2,z3,...,zn}$	-	Partial coefficient of *y* and *x* after when z_2 until z_n are controlled.
$r_{yz1\|z2,z3,...,zn}$	-	Correlation coefficient between *y* and z_1 when z_2 until z_n are controlled.
$r_{xz1\|z2,z3,...,zn}$	-	Correlation coefficient between *x* and z_1 when z_2 until z_n are controlled.
$\mathbf{R}$	-	Optimal rotation matrix.
s	-	Optimal dilation value.
$\mathbf{S}$	-	Unknown independent variables matrix in the ICA algorithms.
S_{m1}	-	Statistic 1 for magnitude of deviation

S_{m2}	-	Statistic 2 for magnitude of deviation
S_r	-	Statistic for relationship
SPE_α	-	α confidence limits for SPE parameter.
SPE_i	-	The contribution of the ith variable to SPE.
tr	-	Trace of matrix.
$\mathbf{t}$	-	Optimal dilation vector.
$\mathbf{t_1}$	-	First PLS scores
T^2	-	T squared value.
T_i^2	-	T squared value at sample i.
T_α	-	α confidence limits for T^2 parameter.
$v_{m,m}$	-	A particular element in eigenvector matrix at row *'m'* and column *'m'*.
$\mathbf{u}$	-	Input vector in the CVA algorithms.
$\mathbf{U}$	-	Eigenvector matrix of major product moment.
V	-	Process variables.
$\mathbf{V}$	-	Eigenvector matrix for PCA.
$\mathbf{V_{PA}}$	-	Eigenvalues matrix in the PA algorithms.
$\mathbf{w}$	-	Noise vector for **x** prediction in the CVA algorithms.
$\mathbf{w_1}$	-	First PLS loadings.
$\mathbf{W}$	-	Unknown transformation matrix in the ICA algorithms.
$x_{n,m}$	-	A particular element at row *'n'* and column *'m'* in a matrix.
$x_{j,i}$	-	Original measurement for variable *'i'* at sample *'j'*.
$\breve{x}_{j,i}$	-	Standardized data for variable *'i'* at sample *'j'*.
$\bar{x}_i$	-	Mean for variable *'i'*.
$\mathbf{x}$	-	System state vector in CVA algorithms.
X_i	-	Contribution of the ith variable to sCMDS/CMDS-PA/CMDS-dPA statistics
$\mathbf{X}$	-	Original data matrix.
$\mathbf{X_E}$	-	Reconstructed multivariate scores by using CMDS.
$\mathbf{X_{mod}}$	-	Modified samples.
$\mathbf{X_{MWOS\text{-}NOC}}$	-	NOC data in terms of MWOS samples.
$\mathbf{X_{NOC1}}$	-	The first set of NOC data used in CMDS algorithms.

$\mathbf{X}_{\mathbf{NOC2}}$ - The second set of NOC data used in CMDS algorithms.

$\breve{\mathbf{X}}$ - Standardized data matrix.

$\hat{\mathbf{X}}$ - Prediction of original data based on PCs model.

$\mathbf{X}_2$ - First residual in PLS algorithms.

y_{dPA} - MDS scores developed through CMDS-dPA algorithms.

y_{PA} - MDS scores developed through CMDS-PA algorithms.

$y_{PA-NOC1}$ - NOC1 scores developed through CMDS-PA algorithms.

y_{sCMDS} - MDS scores developed through sCMDS algoritthms.

y_{NOC1} - NOC1 scores developed through CMDS algorithms.

$\mathbf{y}$ - Output vector in in CVA algorithms.

$\mathbf{Y}_{\mathbf{MWOS\text{-}NOC}}$ - Moving window NOC scores.

$\mathbf{Y}_{\mathbf{NOC}}$ - CMDS-NOC scores.

$\mathbf{Y}_{\mathbf{NOC1}}$ - CMDS-NOC2 scores.

$\mathbf{Y}_{\mathbf{NOC2}}$ - CMDS-NOC2 scores.

$\mathbf{Y}_{\mathbf{dPA}}$ - MDS scores developed through CMDS-dPA algorithms.

$\mathbf{Y}_{\mathbf{sCMDS}}$ - MDS scores developed through sCMDS algorithms.

$\mathbf{Y}_{\mathbf{PA\text{-}NOC}}$ - MDS scores developed through CMDS-PA algorithms.

$\mathbf{Y}_{\mathbf{PLS}}$ - Product quality data matrix in PLS algorithms.

z_{α} - Standard normal deviate corresponding to the upper (1-α) percentile.

$\mathbf{z}_i$ - PCs scores in factor analysis algorithms.

$\mathbf{Z}_{\mathbf{O}}$ - PC scores in terms of observations.

$\mathbf{Z}_{\mathbf{V}}$ - PC scores in terms of variables.

α - Level of confidence limits.

$\delta_{m,m}$ - A particular element in dissimilarity matrix at row 'm' and column 'm'.

$\delta_{NOC1(ij)}$ - Dissimilarity element of sCMDS-NOC1 scores.

$\delta_{PA-NOC1(ij)}$ - Dissimilarity element of CMDS-PA-NOC1 scores.

λ_m - A particular element of eigenvalues for PC 'm'.

ρ - Eigenvalue ratio in CMDS algorithms.

σ_i - Standard deviation for variable 'i'.

ν - Variances of MDS ststistics.

χ	-	Chi-squared distribution.
$\mathbf{\Delta}_{m \times m}$	-	Dissimilarity matrix with size ‘m’ by ‘m’.
$\mathbf{\Delta}^2$	-	Squared dissimilarity matrix.
$\mathbf{\Lambda}$	-	Eigenvalues matrix for PCA.

CHAPTER 1

INTRODUCTION

1.1 Background

The ultimate aim of any production system, in the context of chemical and process industries, is to produce the maximum amount of consistently high quality products as per requested and specified by the customers. This is regarded as highly challenging due to the nature of the processes that always change over time and are also affected by various factors such as variations of raw materials as well as operating conditions, the presence of disturbances and also modification in the process technologies. On top of those, the influence of the surrounding factors including changing in market demands, the environmental impacts, restructuring of the workforces and the unpredictable revolution in the management policies, may also to certain extent affect the product quality as well as the productivity of the production system. In any of the situations, one of the main critical problems is to promptly detect the occurrence of faulty or abnormal operating conditions in the routine process operation and subsequently remove them.

Such issues can be addressed quite effectively by the use of process monitoring techniques. Currently, there are various types of process monitoring systems. However, multivariate statistical process monitoring (MSPM) can be considered as the most practical method for monitoring complicated and large scale industrial processes (Chiang et al., 2001). MSPM mainly addresses multivariate processes with highly correlated process variables where the multivariate correlation among process variables should be essentially considered and cannot be simply ignored during the monitoring operation. Traditionally, statistical process control (SPC) has been used as the basic solution for monitoring industrial production processes. However, this scheme has been criticized, particularly in dealing with multivariate processes, based on various technical aspects as listed as follows:

i. SPC normally monitors the process based on charting only a small number of variables (mostly the variables concerning the final product quality) and examining them only one at a time (Bersimis et al., 2006; Nomikos and MacGregor, 1995).

ii. As a result of (i), analysis and justification cannot be performed accurately as the observed trends are actually reflecting on the correlations among the main variables and not really on the manifestation of individual drive (MacGregor and Kourti, 1995; Raich and Cinar, 1996).

iii. The potentials of false alarms may be multiplied as the variables being monitored increase when applied to multivariate processes (Kano et al., 2002; Mason and Young, 2002).

iv. There are many individual control charts need to be monitored concurrently and this is difficult for process operators to handle (Bersimis et al., 2006).

MSPM, which can be viewed as the upgraded version of SPC, on the other hand, has two significant advantages in relation to the same context as follows:

i. It is a data-driven method (Chiang et al., 2001) and can handle enormous amount of process data (Yoon and MacGregor, 2000; Zhao et al., 2004), which is typical in modern chemical industry as a result of advanced instrumentation and data acquisition techniques. Thus, the excess storage of process data can be beneficially utilized in this regard.

ii. It belongs to the area of process chemometrics (Qin, 2003) where process behaviour is modelled by transforming the high-dimensional data into a lower dimensional space. In effect, some useful parameters (normally in terms of statistical indexes) can be computed subsequently. Therefore, the tasks of monitoring can be made much simpler, where only a small number of control charts are required to be monitored.

In general, the MSPM method normally utilizes two types of monitoring statistics namely Hotelling's T^2 and squared prediction errors (SPE) (MacGregor and Kourti, 1995; Wise and Gallagher, 1996; Martin et al., 1996; Raich and Cinar, 1996). These parameters are derived directly from the multivariate scores in terms of observation-sample configurations. T^2 reflects the magnitude of the deviation of the individual sample from a specified mean, whereas SPE describes the variation in terms of correlations among the monitored process variables (Qin, 2003). The scores are usually developed on the ground of variance-covariance or correlation measure of the standardized multivariate data. Usually, principal component analysis (PCA) is the most widely applied multivariate tool in this regard due to perhaps its simplicity in terms of principles and computation. The main idea of this multivariate technique is simply to transform the original structure of the multivariate data into a few linear combinations (new variables) which mostly capture the original data variations and more importantly these linear combinations are uncorrelated with each other (Jackson, 1991; Gnanadesikan, 1997; Jolliffe, 2002). Thus, it provides the new basis vectors of transformation. In addition, a set of warning and control limits are identified for both statistics and a fault situation will be immediately signalled and declared if any of the statistics violate the corresponding control limits. The Shewhart-type control chart is usually employed in providing the visualisation of the progression of the monitoring statistics (Bersimis, 2007).

The whole framework of MSPM is fundamentally guided as well as complies with the core characteristics of any process monitoring systems as suggested by Jackson (1991) as stated as follows (in no particular order and these criteria will be known afterwards as JPMC – Jackson's Process Monitoring Criteria):

i. A solution should be available in answering this question - 'is the process really in control?'.

ii. Statistical limits of 'Type I' error should be specified reflecting the process environment.

iii. Relationships among the variables should be considered in the monitoring operation.

iv. The source of the problems should be examined in relation to the out-of-control control situation.

Out of these four criteria, the third should be considered as primary as it relates to the emergence of multivariate analysis in the realm of process monitoring. Besides, this is also strongly connected to the problematic situation of implementing the conventional SPC as described previously. In assessing the matter more seriously, two critical issues need to be clarified. The first pertaining to the accuracy of the multivariate model in emulating the true behaviour of the original multivariate data, whereas the second relates to the capability of that particular multivariate model in detecting various types of process malfunctions.

1.2 Motivation

With regard to the first matter, linear PCA is sometimes improperly used especially in modelling highly non-linear processes as a high number of principal components (PCs) is always required technically in order to obtain high degree of transformed variances from the original data (Dong and McAvoy, 1996). Two options are normally available in confronting with this particular issue.

Firstly, non-linear PCA (Dong and McAvoy, 1996) was proposed where less dimensionality can be achieved and eventually feasible for utilization in nonlinear process monitoring. The works reported in (Martin et al., 1996; Zhang et al., 1997; Doymaz et al., 2001; Lopes and Menezes, 2004) constitute the integration technique between auto-associative neural network and principal curve, however, their development is very computationally demanding. In particular, the method involves a huge amount of computation for developing a nonlinear PCA model as well as introduces complications in deciding how many mapping and de-mapping layer nodes are necessary in building up an optimized multivariate model in the form of an auto-associative neural network. Choi et al. (2004) and Lee et al. (2004) have also criticized the approach and they have come out with another version of nonlinear PCA - kernel PCA (KPCA) originally proposed by Schölkopf et al. (1998). Unlike the previous method, which specifically emphasize on the nonlinear optimization, the KPCA scheme focuses on transforming the original data space into a new space (named as the 'feature space') that comprises of higher dimensionalities. The main benefit is that the original nonlinear behaviour can be mapped into the feature space and then analysed through linear correlation (through a specified means of kernel function), and as a result, linear PCA can be effectively executed for monitoring. Even though the computation is technically less complicated, but the implementation is still involved with high dimensionality models, thus the original problem cannot be solved efficiently.

The second option is to permanently engage with the high dimensionality selection. This means, a large number of PCs are selected for application. Nonetheless, this has to involve with cumbersome tasks and is time consuming because there are too many plots that must be analyzed simultaneously (even though this can be resolved with the advanced computing). Besides, the benefit of using PCA for data compression is fundamentally lost (Zhang et al., 1997).

The second issue (relates to the last paragraph in Section 1.1), however, draws the attention in assessing whether a developed monitoring model can reliably give overall excellent process monitoring performance. Despite the fact that non-linear PCA method may tend to produce the correct fault detection outcomes, nonetheless, fault identification may lead to false conclusions as the original non-linear behaviour is still inherited within the modified non-linear PC models. Such inconsistency has been explained by Doymaz et al. (2001) when differentiating the faults that are contributed by sensor failures from those of abnormal processes.

In short, using a linear approach for modelling extremely non-linear multivariate processes is inappropriate as explained in the previous descriptions. Introducing non-linear PCA might seem to be a good answer, but it by no means changes the nature of non-linearity implication of the original multivariate data. Hence, these arguments are the core foundation of this work in order to identify other alternatives which are not just suitable in providing a radical solution from a different perspective (in relation to other multivariate techniques), especially on summarizing the associative data for the use in monitoring applications.

In the effort to understand more deeply towards the issues, the main criteria of multivariate analysis (MVA), which includes PCA, should be conceptually clarified in terms of its functionality. From a generic outlook, the subject interest of MVA is either on analyzing the association among the variables alone or evaluation on the combination between variable-object behaviour (Green and Carroll, 1976). In another viewpoint, Dillon and Goldstein (1984) have simplified the technique into two subdivision methods – dependence and interdependence, where their respective mechanisms are absolutely similar to the previous descriptions. Anyway, both groups have agreed that PCA belongs to the second category, where the aim is to construct a set of reduced dimensional composites which conceptually represents a significant amount of the original multivariate data information.

Under the same category, multidimensional scaling (MDS) is another multivariate technique which can also be utilized in the same regard as PCA. Nevertheless, this technique applies inter-dissimilarity settings of the original multivariate data in developing the multivariate scores in the reduced dimensional space. In other words, the scores are constructed such that the dissimilarity measures of the scores must somewhat be mapped to the dissimilarity measures of the original data. In particular, the points (scores) in the lower dimensional space are arranged in such a way that their distances correspond to the pre-defined dissimilarity measures in the original data set (Takane, 2003). Kruskal and Wish (1978) and Torgerson (1967) stated that the main purpose of any MDS algorithms is to measure how fit the reconstructed multivariate scores configuration matches as closely as possible to the original dissimilarity scales.

MDS has a very strong connection to PCA in two aspects. Firstly, Cox and Cox (1994) outlined the procedures of reproducing the eigen basic structure of dissimilarity scales that is originated from the basis of minor product moment of the original data (which is basically applied for PCA). Secondly, Cox and Cox (1994) and Cox (2001) have proved that both techniques share the same observation-sample score configurations whenever Euclidean-distance is used as the dissimilarity measure. This shows that MDS can contribute in the sense that, firstly the non-linearity issue can be naturally incorporated in the MDS scores through inter-distance measures. Secondly, fault detection can perhaps be operated effectively as well as efficiently, in relatively less dimensionality compared to linear PCA because MDS never use the concept of variance transformation but focusing on dissimilarity scale transformation. This basically means that the magnitude of variances will be identified or formulated based on the multivariate score behaviours instead of adopted from the original data. Therefore, this study believes that MDS has the full potentials based on its unique capacities and is the closest candidate with regard to the main issue. However, the attempt of embedding the technique into the general MSPM framework is not an easy task and has to face four great challenges. The first regards to the multivariate scores development, where MDS essentially uses the inter-dissimilarity measures (usually in terms of distance) in describing the association among the interested objects. One of the main criteria of any

MSPM procedures, suggested by Jackson (1991), is to fundamentally consider the relationships among the variables and not on the observation samples. Thus, in complying strictly according to the rule, the final score coordination will be in the form of variable configuration and not by way of sample distribution as normally obtained through PCA. As a result, the typical T^2 and SPE statistics cannot be utilized in this context. Therefore, the second challenge is to properly introduce different types of monitoring statistics which conceptually correspond to the original definition of the traditional T^2 and SPE as a result of radical change in the scores' basis. Then, the third main task is to propose various appropriate projection mechanisms, as CMDS never utilize any kind of loading factors for the development of the new scores. Lastly, the final critical assignment is to evaluate the performance of the proposed MDS-based monitoring outcomes in relative to the linear PCA based MSPM results in order to justify the beneficial of using the proposed methodology. More specifically, this study attempts to investigate the following critical questions:

i. What is (are) the most appropriate distance(s) to be used for describing the dissimilarity among the variables under the proposed MDS based monitoring scheme? What is/are the main reason(s) behind of the findings?

ii. What kinds of monitoring statistics or parameters that can be used in the proposed monitoring procedures, which are similar in concept to the traditional T^2 and SPE? How do they work?

iii. How does the MDS-based system deal with on-line monitoring? What is (are) the most optimized condition(s) which can produce improved monitoring performance, while sustaining a good representation of normal operating condition (NOC) behaviour?

iv. How should the monitoring performance of the proposed MDS based monitoring schemes be assessed? What is (are) the generic outcome(s) of the comparative evaluation between the MDS-based and linear PCA-based monitoring results, particularly using relatively lower dimensional models?

1.3 Aim and Objectives

The main aim of this research is to develop a comprehensive MDS-based process monitoring system that works systematically according to the general framework of MSPM as well as acts in accordance with the basic monitoring characteristics as suggested by Jackson (1991). In order to achieve this aim as well as to conduct a thorough evaluation on the proposed system, the following list shows the three primary objectives (frameworks):

i. Framework I: To develop a basic generic procedure of MDS-based process monitoring system that utilizes the standard Classical Multidimensional Scaling (CMDS) algorithms for the construction of the NOC model, control limits and also the new sample scores. This framework is considered as the simplest methodology in this study as it merely uses the established CMDS procedures in developing the scores. It is expected that whenever a fault occurs in the process, the resulting scores in the reduced dimensional space will move away from the normal cluster. Fault detection is carried out based on the changes of the resulting scores in the reduced dimensional space.

ii. Framework II: To develop an integrated generic procedure of MDS-based process monitoring system that merges between CMDS and Procrustes Analysis (PA) algorithms in constructing the NOC model, control limits and also projecting the new sample scores. This particular framework can be perceived as an improved technique upon the first framework because it creates a set of PA transformation factors (emulating the concept of loading functions in PCA) in projecting the new sample scores. In other words, it standardizes the score development procedures through the mapping functions of PA.

iii. Framework III: To develop enhanced procedure of MDS-based process monitoring system that combines CMDS and PA algorithms in modelling the NOC behaviour as well as control limits, but more importantly it dynamically projects the sample scores using different sets of PA transformation factors. By utilizing this particular framework, this study believes that the reconstructed scores can be configured more accurately compared to framework II. This will then perhaps, contribute higher degree of robust process monitoring performance.

1.4 Scopes

The scopes of this study are summarised as follows:

i. This study employs two kinds of distances, Euclidean and City-block, in calculating the dissimilarity measures. Typically, MDS utilizes Euclidean space for the construction of the multivariate scores regardless of the distances used. Thus, it is normal to assume that Euclidean distance is more sensible in this situation. Nevertheless, it is also preferable to have another dissimilarity measure, such as City-block, in order to analyze the credibility of the results more critically. Therefore, the idea of employing the City-block distance is such an effort to identify the tendency of changes that other distance measures may contribute compared to the Euclidean distance measure.

ii. This study also introduces two new monitoring statistics, which fundamentally correspond to the original concept of T^2 and SPE. This is very important as this study focuses on the variable score instead of observation-sample structure. In technical terms, the traditional T^2 highlights the weighted distance from a monitored sample to the centre of NOC in the principal component space, while SPE signifies the consistency of the individual scores according to the PCA model. Whenever a fault condition occurred, at least one monitoring statistic will increase in magnitude, but this depends on how sensitive are the chosen PCs in capturing the abnormal trends. The same principle is also applied in the MDS case, but, technically in different practical mechanisms. In the case of the first new monitoring statistic, every score will be either measured according to the centre of the NOC's multivariate score coordinates (origin) or from a specified location that is determined by the NOC model. The second monitoring statistic, on the other hand, connects to the measure of dissimilarity scales between the current samples and original NOC data. It is in this particular step that MDS has an advantage over PCA, whereby MDS is particularly able to reproduce the dissimilarity scales in the reduced dimensional space relative to the original measure. More specifically, the consistency of the variables association can be transferred into the compressed domain, where the analysis can be executed more directly without having to predict back into the original space. Thus, the needs of dealing with more dimensionality can be avoided to certain degree.

iii. This study utilizes the over-lapping moving window schemes (Kano et al., 2000; 2001; 2002) in monitoring continuous process data and this is very vital based on two reasons. Firstly, the distances among process variables under the MDS scheme are calculated based on the number of samples in the NOC data. When calculating the distances among process variables in on-line monitoring, the data dimension should be the same as that in the NOC data. Thus, the over-lapping moving window approach ensures that the distances among process variables can be calculated at every sampling time. The second reason is that the size of moving window structure can be modified as necessarily as required and this initiates various options for the improvement of monitoring performance, particularly in balancing up with the NOC model robustness.

iv. This study examines all the results based on three major performance factors, number of cases detected, fault detection time, and false alarm rates. The descriptions of those terms and how they differentiate with each other will be discussed in detail in subsequent chapters. The emphasis will always be to perform a comparative analysis on the monitoring outcomes between each of the frameworks with the linear PCA based MSPM scheme as well as among of the new methods themselves. This has been done on several fault cases of simulated industrial process (the details are given in Chapter 3). The objective of the whole evaluation is to determine the best possible monitoring solution(s) as well as to identify the impact of using much lower dimensions on the monitoring performance. This is imperative as it highlights the main advantage of using MDS over linear PCA for monitoring the specific continuous multivariate process which tends to be non-linearly behaved.

1.5 Contributions

The main contributions of this study are:

i. The introduction of MDS, or specifically CMDS, as an alternative tool for data dimension reduction technique in the MSPM framework. As a result, a new kind of MSPM framework is eventually obtained and it possesses the following features which are characteristically unique compared to the conventional method:

a. Dissimilarity measure is introduced as the basis in describing the association among of the monitored process variables. The benefit of this method is that it indirectly addresses the issue of using linear methods for modelling non-linear processes.

b. The proposed monitoring system uses variable scores instead of sample scores (as normally obtained through PCA). This main feature enables the monitoring operation as well as its analysis to be performed directly in the reduced dimensional space and not involving the original domain. This will help in the effort of inculcating a relative less dimensional model for an effective monitoring execution.

c. Three types of generic MDS-based MSPM frameworks are proposed and this may initiate other development directions in expanding the process monitoring scopes much wider instead of moving deeper using the PCA approach. Possible further extensions are described in the final chapter under the section of recommendation for further works.

ii. In general, the main strategy, to which the justification process will be made, is by a comparative analysis between the MDS based monitoring system and the linear PCA based monitoring system. Thus, even if the goal and also the objectives of this study have been successfully fulfilled, but this does not prove that the linear PCA is no more applicable. In fact, the main intention of this study is actually trying to promote MDS as an option for improving the linear PCA based monitoring system performance. Hence, this is the main perspective maintained throughout this project especially when discussing the results. In other words, the impact of this study is not to discard the linear PCA from the process monitoring area but rather to strengthen the practicality of MSPM as to be consistently relevant in all situations.

1.5 Thesis Organization

This thesis is divided into seven chapters. Chapter 1 discusses the background of the work which highlights its general motive, importance as well as the main considerations, in which this research should be conducted.

Chapter 2 describes the mathematical theories as well as a critical review of MSPM particularly from the perspective of JPMC and the relations with the motivation of this study. Among others, various extensions of PCA as well as other multivariate statistic techniques on monitoring applications are considered, and are subsequently followed by the description on the fundamentals of MDS. The generic outcome of this particular chapter is to explain the necessity of applying MDS in enhancing the MSPM system, where previous works of MSPM extensions are critically reviewed.

Chapter 3 introduces the case study, which is a continuous stirred tank reactor (CSTR) with recycle through a heat exchanger. In this case study, 11 different types of fault cases will be investigated. The PCA models corresponding to those NOC sets are also presented. Meanwhile, each of these abnormal events is categorized into two groups – abrupt and incipient faults. All of these cases are used to test each of the proposed frameworks.

Chapter 4 focuses on the description of framework I. The corresponding detailed procedures centralised on promoting the variable scores as the basic measure in computing the monitoring statistics are presented. The overall monitoring results are presented and compared with linear PCA based monitoring system. This is to establish a basic perception on the credibility of the assumptions made upon the MSPM applications.

Chapter 5 elaborates thoroughly on the procedures as well as the results of using framework II. The discussion stresses on the extension it has made from framework I, where eventually it changes the way of monitoring statistics is formulated from the initial version. Apart from being compared to the traditional PCA performances, the results of this particular method are also assessed in relative to the outcomes of framework I. This is to understand the implication of the modification, either it alter the original perception (obtained in framework I) or otherwise.

Chapter 6 accentuates on the framework III explanation. The description elaborates the justification of using the dynamic mechanism, but still retaining the steps designed in framework II. While the details might differ slightly from the second framework, the analysis on the results should be conducted comprehensively which include linear-PCA, and both frameworks I and II. This is again to evaluate whether the initiative of injecting the dynamic projection can really create improvement or not on the basic perception obtained from frameworks I and II.

Lastly, Chapter 7 concludes the whole study and presents suggestions for future works. The main consideration of this chapter is to provide answers corresponding to the list of research questions provided in Chapter 1. This has to be performed in light of the result evaluations from those proposed frameworks performances. Some further works are also recommended to expand the capabilities of MDS based monitoring systems proposed in this study.

CHAPTER 2

MULTIVARIATE STATISTICAL PROCESS MONITORING

2.1 Introduction

It was mentioned in the first chapter that MSPM is considered as the best option in monitoring complex as well as considerably large scale industrial systems. There are also other methodologies such as model-based and knowledge-based techniques, but impractical particularly when concerning the huge scale and complexity issues (Chiang et al., 2001). In general, such difficulties are the rigidity, validity as well as difficulty in the development of first principle models, credibility of the process knowledge used as well as spurious decision outcomes, and not to mention complications as well as inflexibility in updating recent information for the improvement of monitoring operation (Venkatasubramanian et al., 2003a; 2003b; 2003c). Nevertheless, these non-statistical process monitoring techniques may undoubtedly become more productive when concerning the diagnostic phase in contradiction to MSPM which is heavily dependent on the credibility of the process history data alone. Venkatasubramanian et al., (2003a; 2003b; 2003c) have also suggested that this can be modified further, perhaps by using a hybrid system that integrates various sets of techniques which works complementary with each other.

This chapter mainly concentrates on the survey of MSPM works and also the extent of the relevancy of those works in relation with this particular study. Besides, this chapter also explains the justifications of using MDS for MSPM applications.

This chapter is divided into five main sections including the first section as the introduction. Section 2.2 discusses the fundamentals of the conventional MSPM methodology with respect to its generic framework and also the corresponding mathematical background. Section 2.3 presents the critical review on the various advanced techniques of industrial-based MSPM applications in the light of Jackson's process monitoring criteria (JPMC) as introduced in the first chapter. Section 2.4 focuses on the implementation of MDS for process monitoring, where the emphasis is on the rationality of applying MDS in the generic MSPM framework. Lastly, Section 2.5 summarizes the chapter.

2.2 Fundamentals of Conventional MSPM System

This section focuses on the general procedures as well as theoretical basis of MSPM techniques. There are also other terminologies such as 'Multivariate Statistical Process Control' (MSPC) (Martin, et al., 1996; Kano et al., 2000; 2001; 2002; Bersimis et. al., 2007) or 'Multivariate Methods for Process Monitoring' (Kourti and MacGregor, 1995) or 'Statistical Process Control' (SPC) in multivariate process (MacGregor and Kourti, 1995) or 'Statistical Process Monitoring' (SPM) (Raich and Cinar, 1996) that have been used to represent the MSPM methodology. In other words, all of these systems denote the same monitoring mechanism that systematically utilizes statistical analysis in capturing the essential process information based on a correlation model from a set of variables of the collected historical normal operational process data (Yoon and MacGregor, 2000). Nevertheless, the depth of the monitoring scopes defined by those works differs from one to another.

According to Chiang et al., (2001), the complete procedures of any process monitoring systems can be generally associated with four main elements, which are fault detection, fault identification, fault diagnosis and process recovery. Fault detection is always the essential and the most basic step (Qin, 2003) and its purpose is to designate the departure of observed samples from an acceptable range using a set of parameters (Himmelblau, 1978). Therefore, fault detection should be accepted as the first step in any of process monitoring system.

With regard to fault identification and fault diagnosis steps, however, the terms are sometimes interchangeably as well as in many cases separated, and thus, a clear standpoint should be described accordingly (it is also understandable that both steps are only executed after a fault has been detected through the fault detection step). In particular, fault identification involves identifying the observed process variables that are most relevant to the fault which is typically identified by using the contribution plot technique, whereas fault diagnosis specifically determines the type of fault which has been significantly (and should be also validated) contribute to the signal (Chiang et al., 2001). Raich and Cinar, (1996) also express the similar idea and they clarified that monitoring mainly concentrates on the detection and identification alone, while diagnosis (sometimes also referred to as isolation) should be performed during the intervention or control phase. MacGregor and Kourti, (1995) as well as Yoon and MacGregor, (2000) have used contribution plots for diagnosis (isolation) operation. This can be achieved by implementing a time series analysis on the contribution plot progression. However, Martin et al,. (1996) argued that the outcomes of the contribution plots do not necessarily indicate the diagnostic solution because the probable cause may be affected by several combination of large (including the largest deviation value) and small magnitude changes. Qin, (2003), on the other hand, has introduced fault reconstruction in the combined phase of fault identification and diagnosis besides of merely using the contribution plot technique. The issue becomes more complicated when considering the process recovery (or control) step.

In conjunction to this, Nomikos and MacGregor, (1995) comment that MSPM is not specifically a cause-and-effect model but rather a general structure of correlation measure of the whole set of variables involved. Therefore, they have suggested using other on-line diagnostic tools in further exploring the nature of the faulty behaviour corresponding to the fault detection outputs. In supporting this argument, recent studies show that MSPM requires more complex solutions such as knowledge-based systems (Norvilas et al., 2000; Undey et al., 2003; Chew et al., 2007) for fault diagnosis.

Thus, this study has summarized the MSPM procedures as shown in Figure 2.1. From Figure 2.1, the generic monitoring framework adopted in this study comprises of two main phases (Mason and Young, 2002; Bersimis et al., 2007), namely off-line modelling and monitoring (phase I) and on-line monitoring (phase II). The final outcomes are the fault detection and identification instead of diagnosis or even control. There are two primary reasons in arriving to this particular decision. Firstly, the focus of the framework should be obliged upon the angle of the problem that this study is trying to address. As far as this matter is concerned, this primarily involves the analysis on the correlations among the variables, which is the third element in JPMC. This factor is very fundamental and strongly relates to the measure of capability of the new fault detection technique as compared to the conventional method. In other words, it is expected that the main obvious impact will be on the fault detection and identification performances rather than diagnosis.

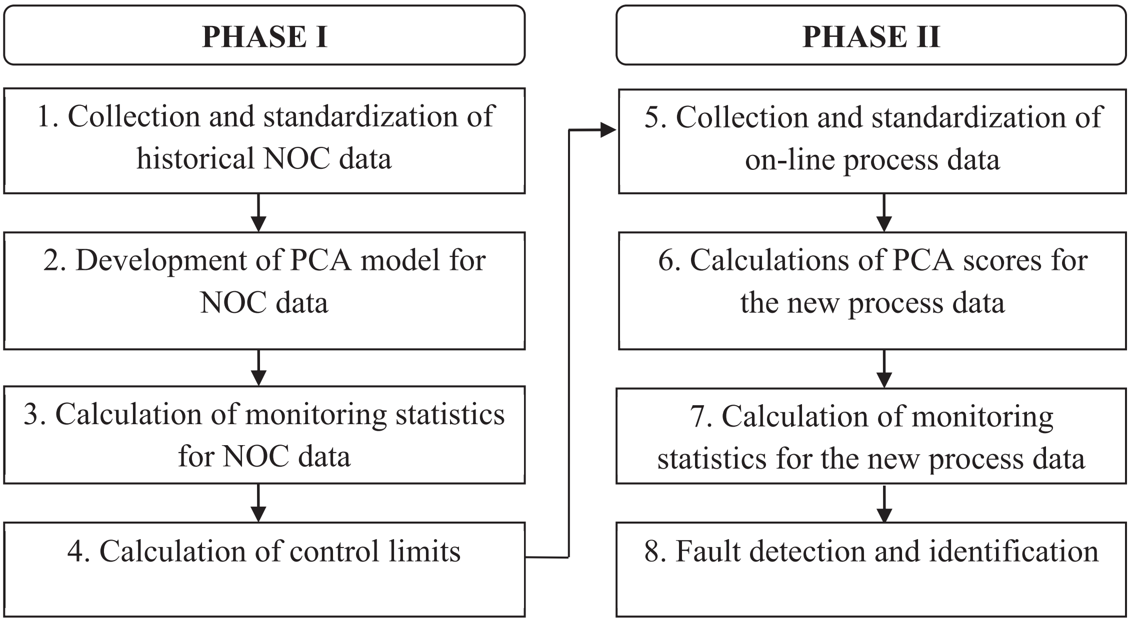

Figure 2.1. Conventional MSPM framework

2.2.1 Phase I: Off-line Modelling and Monitoring

The following discussion is related to steps 1 to 4 in Figure 2.1. Firstly, a set of normal operating condition (NOC) data, **X** of size $n \times m$ (n: number of samples, m: number of variables), are identified off-line (normally steady state) based on the historical process data archive as shown in Equation 2.1.

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{bmatrix} \tag{2.1}$$

where $x_{n,m}$ is a particular element in the NOC data matrix at row n and column m., i.e. variable m measured at sampling time n. Note that this study assumes that standard pre-screening techniques, such as missing data and outlier treatment, have been carried out prior to NOC data determination, and thus. detailed descriptions of data pre-screening techniques are omitted from this thesis.

NOC simply implies that the process is operated at the desired setting condition and produces satisfactory products that meet the specified qualitative as well as quantitative standards (Martin et al., 1996). The data are then standardized to zero mean and unit variance with respect to each of the variables by using Equation (2.2) because PCA results depend on data scales.

$$\breve{x}_{j,i} = \frac{(x_{j,i} - \bar{x}_i)}{\sigma_i} \tag{2.2}$$

Where, $\breve{x}_{j,i}$ = standardized data for variable '*i*' at sample '*j*'.

$x_{j,i}$ = original measurement for variable *'i'* at sample *'j'*.

$\bar{x}_i$ = mean for variable '*i*'.

σ_i = standard deviation for variable '*i*'.

The mean and standard deviation for each of the variables can be obtained respectively through Equation (2.3) and Equation (2.4).

$$\bar{x}_i = \frac{\sum_{j=1}^{n} x_{j,i}}{n}\ ;\ n = \text{number of samples} \tag{2.3}$$

$$\sigma_i = \frac{1}{n-1}\sum_{j=1}^{n}(x_{j,i} - \bar{x}_i)^2\ ;\ n = \text{number of samples} \tag{2.4}$$

As a result, **X** is finally transformed into $\breve{\mathbf{X}}$ (standardized NOC data) after normalization process while the original size is maintained.

In the second step, the development of PCA model for the NOC data requires the establishment of a variance-covariance matrix, $\mathbf{C}_{mxm}$, as indicated in Equation (2.5).

$$\mathbf{C} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,m} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,m} \end{bmatrix} \tag{2.5}$$

Where, c_{mm} = a particular element in the variance-covariance matrix at row *m* and column *m*.

Green and Carroll, (1976) proposed using the sums of squared and cross products (SSCP) matrix in order to obtain **C** through applying minor product moment on $\breve{\mathbf{X}}$ as given in Equation (2.6).

$$\mathbf{C} = \frac{1}{n-1}\breve{\mathbf{X}}^T\breve{\mathbf{X}} \tag{2.6}$$

C is then transformed into a set of basic structures of eigen-based formula as provided in Equation (2.7).

$$\mathbf{C} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T \tag{2.7}$$

Where, **V** = eigenvector matrix of **C** with the size of *m* by *m*.

Λ = eigenvalues matrix (diagonal) of **C** with the size of *m* by *m*.

Finally, the PCA model of $\breve{\mathbf{X}}$ can be simply developed by using Equation (2.8), where **P** is the score matrix and **V** is the loading matrix. The score matrix can be further expanded in details as shown in Equation (2.9).

$$\mathbf{P} = \breve{\mathbf{X}}\mathbf{V} \tag{2.8}$$

$$\mathbf{P} = [\mathbf{p_1} \quad \cdots \quad \mathbf{p_m}] = \begin{bmatrix} \breve{x}_{1,1}v_{1,1} + \cdots + \breve{x}_{1,m}v_{m,1} & \cdots & \breve{x}_{1,1}v_{1,m} + \cdots + \breve{x}_{1,m}v_{m,m} \\ \vdots & \cdots & \vdots \\ \breve{x}_{n,1}v_{1,1} + \cdots + \breve{x}_{n,m}v_{m,1} & \cdots & \breve{x}_{n,1}v_{1,m} + \cdots + \breve{x}_{n,m}v_{m,m} \end{bmatrix} \tag{2.9}$$

From Equation (2.8), **V** provides the linear weighting functions (loading factor) that also reflect the correlation with respect to the original variables. **P** in Equation (2.8), on the other hand, represents the PCs scores matrix and it contains '*n*' number of scores for each of the PCs and structured in terms of linear composites between $\breve{\mathbf{X}}$ and **V** as denoted in Equation (2.9). This particular model is the main justification which highlights the feasibility of using the MSPM-based techniques for monitoring industrial processes, whereby all variables are connected under a single equation to calculate a principal component score regardless of the amount of variables used (as depicted in Equation 2.9). The essential information is originally retained by embedding the original variation in certain degree (number of PCs) into the structure of PCA model. Typically, only a set of the first a few PCs, *a*, are used with $a < m$.

Equation (2.10) presents a measure of data variations captured by the first a principal components (Jolliffe, 2002).

$$k = \frac{\lambda_1 + \lambda_2 + \ldots + \lambda_a}{\lambda_1 + \lambda_2 + \ldots + \lambda_a + \ldots + \lambda_m} \tag{2.10}$$

Where λ_1 to λ_m are the eigenvalues in $\mathbf{\Lambda}$.

The third step basically involves calculation of the Hotelling's T^2 and SPE monitoring statistics. Equation (2.11) shows the mathematical formula for the Hotelling's T^2 statistic (Jackson, 1991; MacGregor and Kourti, 1995; Kourti and MacGregor, 1995; Wise and Gallagher, 1996; Martin et al., 1996; Raich and Cinar, 1996; Mason and Young, 2002)

$$T_i^2 = \sum_{j=1}^{a} \frac{p_{i,j}^2}{\lambda_j} \tag{2.11}$$

Where, $T_i^2 = T^2$ value for sample i.

$p_{i,j} = i^{\text{th}}$ element for principal component j.

λ_j = eigenvalue corresponds to principal component j.

a = number of PCs retained.

Mason and Young, (2002) emphasizes that T^2 is a kind of weighted distance, which is conceptually dissimilar to the original Euclidean distance, where the variation on the samples of the 'new variables' (PCs) is significantly defined by the magnitude of its respective variances-covariances. Besides, Mason and Young, (2002) and Bersimis et al., (2007) have both agreed that the measure is also regarded as Mahalanobis distance (weighted or ratio distance), and according to Hotelling, (1931), it can be strongly connected to the idea of t student statistic (particularly when concerning on the inter-correlation between two groups of samples in its original formula structure). In particular, it indicates the scale of deviation of its individual averaged sample mean with respect to the population mean, where the value will be typically around zero in the normal condition (meaning that it is close to the targeted value).

The SPE statistic, on the other hand, focuses on the consistency of the current variable relationships according to the pre-defined PCA model of the original NOC data (Jackson, 1991; MacGregor and Kourti, 1995; Kourti and MacGregor, 1995; Wise and Gallagher, 1996; Martin et al., 1996; Raich and Cinar, 1996). If the first *a* principal components are retained in the PCA model, then the PCA model predictions of the scaled NOC data are given by Equation (2.12).

$$\hat{\mathbf{X}} = \mathbf{P_a V_a^T} \tag{2.12}$$

Where $\mathbf{P}_a = [\mathbf{p}_1\ \mathbf{p}_2\ \ldots\ \mathbf{p}_a]$ is the score matrix containing the first *a* score vectors and $\mathbf{V}_a = [\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_a]$ is the loading matrix containing the first *a* loading vectors.

The residual matrix, $\tilde{\mathbf{E}}$, between $\hat{\mathbf{X}}$ and $\breve{\mathbf{X}}$ is given by Equation (2.13).

$$\begin{aligned}\tilde{\mathbf{E}} &= \breve{\mathbf{X}} - \hat{\mathbf{X}} \\ &= \breve{\mathbf{X}} - \mathbf{P}_a \mathbf{V_a^T} \\ &= \breve{\mathbf{X}} - \breve{\mathbf{X}}\mathbf{V}_a \mathbf{V}_a^{\mathbf{T}} \\ &= \breve{\mathbf{X}}\left(\mathbf{I} - \mathbf{V}_a \mathbf{V}_a^{\mathbf{T}}\right)\end{aligned} \tag{2.13}$$

Where **I** is the identity matrix.

SPE is then computed as the sum of squared prediction errors as given in Equation (2.14).

$$SPE_i = \tilde{\mathbf{e}}_\mathbf{i} \tilde{\mathbf{e}}_\mathbf{i}^{\,T} \tag{2.14}$$

Where, $\tilde{\mathbf{e}}_\mathbf{i}$ = the *i*th row vector in the residual matrix.

According to Jackson and Mudholkar, (1979), this particular parameter (also referred as *Q*-statistics) should be analyzed critically before assessing the T^2 parameter. This is simply because SPE reflects the assumption made on the variable correlations, thus they commented that 'it is better to test the assumptions first as any invalid assumptions will affect the credibility of the test'. In addition, Qin, (2003) also argued that a radical change detected by T^2 and not captured by SPE is not necessarily due to a fault, but rather a shift in the operating mode. In relation to this, Zhao et al., (2004) produced a comprehensive monitoring procedure that utilized multiple PCA models for managing multiple operating modes. It is generally presumed that SPE is relatively preferable than T^2. However, both monitoring statistics should be used at the same time in order to detect different types of faults.

The final task in phase I deals with developing the control limits for both of the statistics as shown in Equation (2.15) for the T^2 parameter and Equations (2.16) to (2.20) for the SPE parameter (Jackson and Mudholkar, 1979, Jackson, 1991; Wise and Gallagher, 1996; Raich and Cinar, 1996).

$$T_\alpha = \frac{A(n-1)}{(n-A)} F_{A,n-A,\alpha} \tag{2.15}$$

Where, A = number of PCs retained in the PCA model.

n = number of samples.

F = F distributional index with A and n-A degrees of freedom at α confident limit.

α = 95% (warning limit) or 99% (control limit).

$$SPE_\alpha = \theta_1 \left(\frac{z_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} + 1 \right)^{\frac{1}{h_0}} \tag{2.16}$$

$$\theta_1 = \sum_{i=A+1}^{m} \lambda_i \tag{2.17}$$

$$\theta_2 = \sum_{i=A+1}^{m} \lambda_i^2 \tag{2.18}$$

$$\theta_3 = \sum_{i=A+1}^{m} \lambda_i^3 \tag{2.19}$$

$$h_o = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2} \tag{2.20}$$

Where, z_α= standard normal deviate corresponding to the upper (1-α) percentile.

A = number of PCs retained in the PCA model.

m = total number of variables.

α = 95% (warning limit) or 99% (control limit).

Usually, another set of NOC data (sometimes referred as testing set) will be collected and analyzed according to the established monitoring limits through both of the monitoring statistics for assessing its robustness. The details are given in Chapter 3. In this regard, NOC is defined as having the common cause which is entirely due to random noise only (Chiang et al., 2001). All these steps of phase I are also adhered to the generic criteria of monitoring system suggested by Jackson, (1991), which specifically corresponds to the criterion number (i), (ii) and (iii) simultaneously.

2.2.2 Phase II: On-line Monitoring

This section explains the on-line monitoring of the new samples that are collected on-line. Basically, steps 5 to 7 follow similar procedures of steps 1 to 3 in phase I. In step 5, the new samples will be standardized according to Equation (2.21).

$$\breve{x}(new)_{j,i} = \frac{\left(x(new)_{j,i} - \bar{x}_i\right)}{\sigma_i} \tag{2.21}$$

Where, $\breve{x}(new)_{j,i}$ = standardized form of the new data for variable '*i*' at sample '*j*'.

$x(new)_{j,i}$ = original measurement of the new data for variable *'i'* at sample *'j'*.

$\bar{x}_i$ = mean for variable '*i*' of the NOC data.

σ_i = standard deviation for variable '*i*' of the NOC data.

Next in step 6, the new samples (in the standardized form) are projected into multivariate scores by employing Equation (2.22).

$$\mathbf{p_i(new)} = \mathbf{\breve{x}_i}^{\mathbf{T}}\mathbf{V_a} \tag{2.22}$$

Where, $\mathbf{p_i(new)}$ = PCs scores for the new sample *'i'*.

$\mathbf{\breve{x}_i}^{\mathbf{T}}$ = vector of the new sample *'i'*.

$\mathbf{V_a}$ = matrix of eigenvectors based on *'a'* number of selected PCs.

Then, both T^2 and SPE statistics are computed in step 7as denoted in Equations (2.23) and (2.24) respectively.

$$T(new)_i^2 = \sum_{j=1}^{a} \frac{p(new)_{i,j}^2}{\lambda_j} \qquad (2.23)$$

Where, $T(new)_i^2 = T^2$ value for the new sample i.

$p(new)_{i,j}^2 = i^{th}$ element for principal component j of the new scores.

λ_j = eigenvalue corresponds to principal component j.

a = number of PCs retained.

$$SPE\ (new)_i = \mathbf{\tilde{e}(new)_i}\,\mathbf{\tilde{e}(new)_i}^T \qquad (2.24)$$

Where, $\mathbf{\tilde{e}(new)_i}$ = vector for the new sample *'i'*, by which, $\mathbf{\tilde{e}_i} = \mathbf{\breve{x}_i} - \mathbf{\hat{x}_i}$.

Regarding to step 8 (the last step), there are two main operations which have to be conducted separately - fault detection and fault identification. Regarding fault detection, a fault situation is regarded as a result of an occurrence of a special event that is not in conformance to the common cause nature (Chiang et al., 2001). Technically, a fault situation will be declared if either of the monitoring statistics exceeding its respective control limit for a pre-defined successive number of samples. Mason and Young, (2000) used the term 'non-central distribution' in denoting the particular situation with regard to large T^2 values. With respect to abnormal situations flagged up by large SPE, Jackson and Mudholkar, (1979) explained that this abnormality condition can be conceptually regarded as the current variable correlations deviate from that specified by the PCA model (which conceptually represent the NOC behaviour).

Gertler, (1998), has categorized the process faults into 4 main groups.

i. Additive process faults

   External disturbances are affecting the stability of operation normal condition. Great deviation may occur on the variables outputs, which are hardly recognizable through observing on the plant readings. Such events can be related to pipe leakages or abnormal loads.

ii. Multiplicative process faults

Abnormalities in terms of gradual or abrupt changes are observed in some of the plant parameters. The deviations can be spotted by observing both inputs as well as the main output variables. Such faults can be due to deterioration of plant equipment, such as surface contamination, clogging, or the partial or total loss of power.

iii. Sensor faults

This particular type of fault can be resulted from discrepancies between the measured and actual values of individual process variables. Usually, this fault is considered as additive element to the previous main faults as it merely involve with electrical problem which is not directly disturbing the normal operation.

iv. Actuator faults

There are inconsistencies between the input command of an actuator and its actual desired action. This is also can be considered as secondary because it deals with mechanical complexity and not on the fundamental deviation of the normal process.

In addition, the faults can be also existed as an integrated event of one and another. Particularly for faults (iii) and (iv), they can be both appear in the form of multiplicative faults, or in other words, sensor and actuator faults can be monitored through observing the corresponding process parameters. In another instance, Venkatasubramanian et al., (2003), has divided the faults into three main classes, which are gross parameter changes in a model, structural changes and also malfunctions of instruments (sensors and actuators). From their perspective, the first and second relate to additive and multiplicative process faults explained by Gertler, (1998) respectively. Meanwhile, Venkatasubramanian et al., (2003) has combined both categories (iii) and (iv) provided by Gertler, (1998) into one major domain, as well as basically agreed that, these faults can be also occurred as a joined scheme with the previous two types of malfunction conditions. More importantly, Venkatasubramanian et al., (2003), have also provided a generic schematic diagram on connecting both monitoring and conventional feedback control systems, in which, all the signals applied in the four elements of feedback control can be also utilised for process monitoring. Lastly, they have also stated that outside the scopes of these classification, any deviation should be regarded as noise (process or measurement) or suffered from unstructured uncertainties.

Contribution plot is the special tool which is normally used in identifying the variables that are most related to the detected abnormal condition. Technically, the out-of-control statistic is decomposed into components corresponding to different variables and those variables with comparatively large contributions are considered as related to the fault (MacGregor and Kourti, 1995; Yoon and MacGregor, 2000). This operation is easier to be performed based on the SPE parameter, where the contribution for SPE is simply breaking down the summation of SPE as shown in the following equation (Qin, 2003):

$$SPE = \sum_{i=1}^{m} e_i^2 = \sum_{i=1}^{m} SPE_i \qquad (2.25)$$

Where SPE_i is the contribution of the *i*th variable to SPE at a particular sampling time.

Qin, (2003) stated that finding the contribution measure based on T^2 is rather complex as it involves each of the PCs for every variable and may sometimes lead to negative contributions. In responding to the issues, Chiang et al., (2000) proposed to replace the negative values with zero, whereas Westerhuis et al., (2000) proposed utilizing generalized contribution plot to overcome the individual decomposition difficulty. The most common way for the contribution plot for T^2 is to find contributions to the PCs that lead to large T^2 values as shown as follows (Yoon and MacGregor, 2000; Chiang et al., 2000):

$$Cont_{T^2}(x_i) = p_{ij}\breve{x}_i \qquad (2.26)$$

Where, p_{ij} = weight of the *i*th variable in the *j*th PC.

$\breve{x}_i$ = standardized variables.

Contribution plots may not explain explicitly as well as directly the root cause(s) of the problem as the technique itself is not fundamentally a causal-based model but it may considerably assist in narrowing down the investigation task (MacGregor and Kourti, 1995; Yoon and MacGregor, 2000; Qin, 2003). Thus, it all depends on the interpreters (plant operators/process engineers) in executing further analysis on the outcomes or other diagnosis tools should be used in this respect.

2.3 MSPM Issues and Extensions

MSPM has been generally applied in many industrial applications as summarized in Table 2.1, where it also indirectly indicates that MSPM is generally practical for a wide range of industrial monitoring applications.

Table 2.1: Examples of industrial applications of MSPM

Specific Purpose	Industrial Applications
Process monitoring of continuous process	Dairy pasteurization of food process (Kosebalaban and Cinar, 2001); fermentation (Lopes and Menezes, 2004); hot stove system - POSCO (Lee et. al., 2004); pulp and paper Portuguese mill – Portucel (Reis and Saraiva, 2006); iron and steel – Sumitomo Metals Kokura (Kano and Nakagawa, 2008).
Process monitoring of batch/semi-batch process	Polymerisation reactor (Kourti et al., 1995); sugar crystallization (Simoglou, 2005); polymerization – Dupont Co. (Camacho and Pico, 2006); wastewater treatment (Lee et. al., 2006, Aguado et. al., 2007); bioprocess (Cimander et al., 2003, Amigo et. al., 2008);
Loss prevention	Continuous catalyst generator (You, 1998); sensor failure management (Ramaker et. al., 2006); HPLC (Zhu et al., 2007).

Table 2.2: Categories of MSPM issues and extensions reviewed in this study

No.	**Main Categories**	**Sub-categories**
1	Monitoring statistics	1. Integrated parameters 2. Monitoring indexes
2	Monitoring limits	1. Alternative control charts 2. Non-parametric schemes
3	Multivariate techniques	1. Advanced PCA-based techniques 2. Non PCA-based techniques
4	Fault diagnosis	1. Quantitative techniques 2. Qualitative techniques
5	Data formulation	1. Embedding external data 2. Empirical modelling data
6	Automation system	1. Industrial and real-time issues 2. Hybrid system development

However, in the attempt of injecting more attractive and sustainably practical elements into the system, the conventional procedures have been aggressively modified in order to meet the specific various application challenges. Table 2.2 shows issues and extensions of MSPM that are reviewed in this study.

2.3.1 Issue 1: Monitoring Statistics

With regard to the first issue, any MSPM method has to adopt monitoring statistics for indicating the status of the process under monitoring. Typically, a control limit will be set based on the collected NOC data and once this control limit is violated by the monitoring statistic, a fault situation is detected. Yue and Qin, (2001) proposed a single parameter which combines both T^2 and SPE together. The new single index was used for fault reconstruction in order to improve the fault identification operation. In another approach, Chen et al., (2004) argues that the reduced variable set may not be normally distributed as usually presumed. Therefore, they have proposed using kernel estimation approach for the determination of the sample distribution. Nonetheless, the computation is technically complex and as a result, they proposed a synthesis scheme of T^2 and SPE for monitoring. A rather more complicated solution is given by Qin, (2003), where a unified form of various monitoring statistics including T^2, SPE and Mahalanobis distance is produced. As far as these works are concerned, the methods solely concern with fault detection sensitivity. Nonetheless, the original conceptual definition of T^2 and SPE are still unchanged, and thus no radical modification has been imposed in this respect

Unlike the conventional as well as the integrated statistics of MSPM, Kano et al., (2000; 2001; 2002) proposed using 'Moving-window PCA (MW-PCA)' in detecting faults. They argued that changes in the correlation variable can be detected more effective by monitoring the changes in the PCs loadings rather than scores. Consequently, they have introduced monitoring index in replace of statistics, where the index becomes zero when the current sample is operated normally, otherwise it tends close to one. They demonstrated that this method can only be productive when a high number of PCs are retained in the monitoring model. Therefore, this implies that the original concept of variable association formulation is still maintained as original, thus this can create problem when dealing with highly non-linear processes.

2.3.2 Issue 2: Monitoring Limits

The second issue involves with procedures in specifying the control limits. Mason and Young, (2002) commented that the control limits should be carefully set according to the assumption made on the distributional properties, and that corresponds to those parameters (mean and covariance vectors) are either known or unknown. In particular, the *F* distribution is only applicable when the monitoring involves with estimated parameters of mean and covariance vectors of process variables that follow normal distribution, otherwise other distribution models such as chi-square or beta model should be applied. Nonetheless, this suggestion is only relevant when the T^2 is implemented in its original form (not calculated based on the scores in the latent space). Thus, the previous setting limits are still important in this study.

Beside of using Shewhart-type control chart for multivariate monitoring operation, there are also other important tools such as 'Multivariate Cumulative Sum (CUSUM)' (Woodall and Ncube, 1985) and 'Multivariate Exponential Weight Moving Average (MEWMA)' (Lowry et al., 1992) that have been utilized in improving the monitoring performance. The motive is that the conventional control chart is found relatively insensitive to small and moderate shifts in the mean vector due to limitation of information which is based solely on the current sample analysis (Qin, 2003). In contrary, both MCUSUM and MEWMA take the advantage of the past and current values in determining the status of the process based on a specified frame of time. Thus, these techniques have a major impact on the way the mean is determined and eventually on the control limits. Nevertheless, both of these charts are unsuitable to be used in this study based on two reasons. Firstly, the techniques are fundamentally focusing on the mean vector and not considering on the SPE chart. As this study deals with both statistics concurrently, thus, the improvement on T^2 performance alone may not be crucial as this can be complemented by the use of SPE chart. The second reason is that there are no evidence or works which utilized those charts for SPE progression. This could be due to difficulty in explaining the relationship between the MCUSUM and MEWMA theories and the SPE concept.

Another important area which is popularly explored is the use of non-parametric schemes in estimating the probability distribution function (PDF), where one of the major works has been on kernel-based distribution estimation (KDE) (Martin et al., 1996). The conventional Gaussian distribution approach totally depends on the pre-determined parameters, such as mean and covariance, in defining the underlying function. On the other hand, KDE defines the PDF based on the summation in terms of kernel function that represented in the form of 'bumps' or 'humps', which are centred on each data point. Nonetheless, this method does not seriously consider the dependability on the linearity assumption that is used for the variable correlation. Thus, this may tend to involve a high number of PCs when the relationship among process variables is highly-nonlinear. Besides, the technique has only proven effective when it involves small scale processes (Martin et al., 1996).

2.3.3 Issue 3: Multivariate Techniques

The third issue corresponds to the application of multivariate technique and has been perceived as the key element that particularly relates to the contribution of this study. Typically, it is the behaviour of the variables (not on observation correlations or other possible means) of the process under investigation which should be essentially observed, as well as assessed when some changes happened or disturbed in the process, especially in bringing the process back to the normal region. As far as advanced MSPM is concerned, there are two main directions where all the developments can be categorized into.

The first regards to modification or improvement that is based on PCA technique and the others correspond to the utilization of non-PCA techniques. The advanced applications that use the PCA method should perhaps be perceived as the most popularly explored as shown in Table 2.3. In general, Table 2.3 indicates that the method of PCA itself is technically flexible and can be effectively adapted in various situations. There are two kinds of generic applications with regard to PCA extensions that have been made and those include modification within the technique itself (issues number 1 to 6) and the other involves integration with other schemes (issue number 7). Nonetheless, despite being very intensively explored and extended, these advanced works never significantly change the original interpretation on the variable correlation. In contradiction to the goal and studies of non-linear techniques (which have been addressed in Chapter 1), these works still assume that the variables are associated linearly, and as a result, they have to use a high number of PCs for representing sufficient amount of data variations. Therefore, this study recognizes this direction as corresponding to the technical impact and merely involves with secondary issues rather than primary (based on the scopes specified in this study).

Table 2.3 Extensions of PCA-based monitoring techniques

Issues	Solution (s)	Summary Descriptions	References
1. Active changes in operating modes.	Adaptive PCA/MSPC	A weighted mean and covariance matrices are updated and embedded into the current PCA model. Thus, a new improved PCA model is produced which reflects the current operating mode.	Choi et al., (2006)
	Multi-recipes/products PCA	The pooled sample variance-covariance matrix of the individual product is used to estimate the principal component loadings of the multi-group model.	Lane et al., (2001)
2. Handling large scale processes.	Multi-block PCA	The process variables are divided into several blocks with respect to the number of unit operations involved for easier modelling and interpretation.	MacGregor et al., (1994).
3. Dynamic monitoring.	Dynamic PCA	This dynamic approach constructs a time series model from the data, where the detection sensitivity can be achieved especially when involving correlation model inconsistency.	Ku et al., (1995)
4. Handling data that subject to abnormality occurrence at different localization.	Multi-scale PCA	Separate PCA models are determined at different scales for increasing the fault detection sensitivity.	Bakshi, (1998)
5. Facilitating for batch process monitoring	Multiway / unfolding-PCA	A three dimensional array data comprises of process variables, batch numbers and observation samples of each batch are transformed into 2-dimensional structure for the feasibility use of PCA.	Nomikos and MacGregor, (1995), Nomikos, (1996), Lee, et al., (2004a)
6. Multi-phased dynamic monitoring	Multi-phase PCA	Extended version of multi-way PCA where dynamic element is introduced in the monitoring scheme.	Chamaco and Pico, (2006), Doan and Srinivasan, (2008)
7. Handling autocorrelation data.	ARMA filters + PCA	ARMA filter is used to remove the auto-correlation from the monitored variables to avoid the production of false alarms.	Kruger et al., (2004)

In another angle of advancement, the works suggest that the old style of linear technique should be reviewed and sometimes challenged in order to spark improvement in the process monitoring performance. Therefore, the impact of these changes should be regarded as fundamental as well as significant with respect to the particular criterion discussed in this sub-section because it initiates conceptual deviation from the conventional perspective. In general, these works adopt different kinds of multivariate and non-multivariate techniques instead of PCA for summarizing and comprising the multivariate data.

Partial least square (PLS) is perhaps the main competitor of PCA with regard to its popularity in the area of MSPM application. Among others, the original works have been proposed by Nomikos and MacGregor, (1995), as well as Kourti et al., (1995), for batch process monitoring using multi-way PLS, whereas Kourti and MacGregor, (1995) proposed using PLS for both continuous and batch processes. Various extensions that focus on industrial applications also have been developed, for instance, Zhao et al., (2006) have utilised multiple PLS models with principal angles in coping with multiple operating modes, Lee at al., (2006) introduced PLS for biological wastewater treatment plant with non-linear application and Gunther et al., (2009) applied evolving PLS in industrial fed-batch cell culture. In a slightly different form compared to PCA, PLS requires two sets of data, which are a set of variables obtained from the on-line measurements, (**X**) and the other holds a set of quality variables collected either off-line or on-line, ($\mathbf{Y_{PLS}}$). The aim is to capture the variation in **X** which is most predictive of the product quality data **Y**, based on the sample covariance matrix of $\left(\mathbf{X}^T\mathbf{Y_{PLS}}\right)(\mathbf{Y_{PLS}}^T\mathbf{X})$ as shown in the following procedures:

i. First PLS loading = $\mathbf{w}_1$= first eigenvector of $\mathbf{X}^T\mathbf{Y_{PLS}}\mathbf{Y_{PLS}^T}\mathbf{X}$. (2.27)

ii. First PLS scores = $\mathbf{t}_1 = \mathbf{X}\mathbf{w}_1$ (2.28)

iii. First regression factor = $\mathbf{p}_1 = \mathbf{X}\mathbf{t}_1/\mathbf{t}_1^T\mathbf{t}_1$ (2.29)

iv. First residual = $\mathbf{X}_2 = \mathbf{X} - \mathbf{t}_1\mathbf{p}_1^T$ (2.30)

v. Second PLS loading = first eigenvector of $\mathbf{X}_2^T\mathbf{Y_{PLS}}\mathbf{Y_{PLS}^T}\mathbf{X}_2$ (2.31)

vi. Second PLS scores = $\mathbf{t}_2 = \mathbf{X}\mathbf{w}_2$ (2.32)

According to MacGregor and Kourti, (1995), the scores are linear combination of **X**-variables that maximizes the covariance between it and the **Y** space, whereby, the scores are orthogonal with the associated loading vectors. Conceptually, PLS is similar to PCA except that it simultaneously reduces the dimensions of the $\mathbf{X}$ and $\mathbf{Y_{PLS}}$ spaces. This technique is suitably applied in the situations where one may have information on both of the process variables, the input variables $\mathbf{X}$, and the product quality variables, $\mathbf{Y_{PLS}}$, at the same time. In this case, the PLS methods can be used to make relationships between $\mathbf{Y_{PLS}}$ and $\mathbf{X}$. Nonetheless, the mechanism of separating two sets of basic data tends to be problematic when considering the continuous process. This is because continuous process monitoring system treats all the variables equally as well as simultaneously.

Another important technique is called factor analysis (FCA). FCA generally has the similar aims to PCA, except that it has an extension form of structure by way of rescaling on the original PCA model as shown as follows:

$$\mathbf{X}_1 = \mathbf{a}_{11}\mathbf{F}_1 + \mathbf{a}_{12}\mathbf{F}_2 + \cdots + \mathbf{a}_{1m}\mathbf{F}_m + \mathbf{e}_1$$

$$\vdots$$

$$\mathbf{X}_p = \mathbf{a}_{p1}\mathbf{F}_1 + \mathbf{a}_{p2}\mathbf{F}_2 + \cdots + \mathbf{a}_{pm}\mathbf{F}_m + \mathbf{e}_p \qquad \text{... (2.33)}$$

where, $\mathbf{a}_{ij} = \sqrt{\lambda_i}\mathbf{b}_{ij}$; $\mathbf{b}_{ij}$= principal component loadings; $\mathbf{F}_i = \frac{\mathbf{z}_i}{\sqrt{\lambda_i}}$; $\mathbf{z}_i$= principal component scores.

By having such structures, the new uncorrelated scores will have unit variances. Amigo et al., (2008) has used FCA specifically under a model called as Parallel Factor Analysis (PARAFAC) to be inculcated in the MSPC framework as a monitoring and real-time control tool for bioprocesses. Taking advantages of the mathematical properties of PARAFAC, batches of a bioprocess measured under normal operating conditions were used to develop a calibration models in a real time manner. Therefore, the main advantage of PARAFAC is that the true underlying phenomena can be obtained from the model scores without depending on the off-line model.

Canonical variate analysis (CVA) is another technique that is similar in concept to PLS. Thus, two sets of data are generally required and the goal is to capture the linear combinations that signify maximum correlation between these matrices. Simoglou et al., (2002) have proposed using CVA technique and state space models for process monitoring combining past measurements and current/future values. Thus, their concern is on the dynamic behaviour of the system (by way of state space modelling) rather than minimizing the chosen number of multivariate model dimensions. The concept of state space modelling is based on describing a system in terms of k first-order difference equations that are combined into a first-order vector–matrix difference equation as shown as follows:

$$\mathbf{x}(t+1)=\mathbf{F}\mathbf{x}(t)+\mathbf{G}\mathbf{u}(t)+\mathbf{w}(t) \tag{2.34}$$

$$\mathbf{y}(t)=\mathbf{H}\mathbf{x}(t)+\mathbf{A}\mathbf{u}(t)+\mathbf{B}\mathbf{w}(t)+\mathbf{e}(t) \tag{2.35}$$

where **x**, **u** and **y** are the system states, inputs and outputs respectively, while **F**, **G**, **H** and **A** are their corresponding transformation matrices. The terms **w** and **e** denote noise processes that are each assumed to be independent and identically distributed.

In conjunction to this, the study proposed by Treasure et al., (2004) on applying subspace identification in describing the dynamic states of process variables for monitoring also cannot be applied in this study, because the scope is different in context. Perhaps, their approaches can be fully utilized when the fundamentals of this new system have been thoroughly developed and analyzed (this is the main focus concerned in this work).

Unlike PCA, PLS, FCA and CVA, the main function of Independent Component Analysis (ICA), on the other hand, is to decompose the multivariate data into independent sub-components which are presumably to be individually non-Gaussian distributed. The assumption made by ICA is that every observed data, **X** can be expressible in terms of linear combinations between the unknown independent variables and unknown transformation matrices as shown as follows:

$$\mathbf{X} = \mathbf{W}\mathbf{S} + \mathbf{E}, \tag{2.36}$$

Where, **W** = transformation matrix,

S = hidden space matrix

E = residual matrix

Therefore, the objective is to find these latent matrices that correspond to the original value considered. Hence, ICA has a major implication on the interpretation of the variable correlation assumption, and that is, it radically de-correlates the original linear model. A number of works in process monitoring have been developed with regard to this technique such as Lee et al., (2004b), Lee et al., (2004c), Yoo et al., (2004) and Kano et al., (2004).

Even though it was claimed by those works that ICA can give superior performance over PCA, however, Lee et al., (2004b) accepted the fact that this technique suffers some technical difficulties, where there is no standardized method which can be used confidently for conducting the component selection (in contradiction to PCA, where the component ranking can be easily obtained, by relating to the descending order of eigenvalue magnitudes). Thus, this issue may introduce more complexity, especially when concerning on the effort of minimizing the number of components chosen. Hence, similar to CVA, this technique is actually relevant for consideration but it should be performed perhaps on the next advanced level of monitoring.

In contrary to the previous methods, the purpose of Partial Correlation Analysis (PCarrA) technique is to determine the correlation between two variables while keeping the other variable constant. Partial correlation is the correlation of two variables while rigidly controlling another or more other variables. DeVor et al., (1992) has explained that the partial correlation presents the incremental predictive effect of one process variable based from the collective effect of all others, while the main objective is always to find a set of variable combination that produce the largest incremental predictive power.

If the two variables of interest are y and x and the control variables are z_1, z_2 ... z_n, then the corresponding partial correlation coefficient is $r_{yx|z1,z2...,zn}$ as denoted in Equation (2.37).

$$r_{yx|z1,z2,z3,\ldots,zn} = \frac{r_{yx|z2,z3,\ldots,zn} - r_{yz1|z2,z3,\ldots,zn}\, r_{xz1|z2,z3,\ldots,zn}}{\sqrt{1-r^2_{yz1|z2,z3,\ldots,zn}}\sqrt{1-r^2_{xz1|z2,z3,\ldots,zn}}} \tag{2.37}$$

Where, $r_{yx|z1,z2,z3,\ldots,zn}$ = partial coefficient of y and x after when z_1 until z_n are controlled.

$r_{yx|z2,z3,\ldots,zn}$ = partial coefficient of y and x after when z_2 until z_n are controlled.

$r_{yz1|z2,z3,\ldots,zn}$ = correlation coefficient between y and z_1 when z_2 until z_n are controlled.

$r_{xz1|z2,z3,\ldots,zn}$ = correlation coefficient between x and z_1 when z_2 until z_n are controlled.

From Equation (2.37), this correlation is identified by separating process variables into a number of subgroup, in which one or more variables are held constant before determining the correlation among the other variables.

PCorrA has been applied by Ibrahim, (1997) for developing Active SPC technique. The method is similar to those applications used in the feedforward Automatic Process Control (APC) and it was applied to a nonlinear simulated Continuous Stirred Tank Reactor (CSTR).

In another interesting direction, Albazzaz et al., (2005) have explored the practicality of using Multidimensional Visualization (MDV) for process monitoring. One of the obvious advantages of MDV against the conventional PCA-based system is that it breaks the limitation of dimensional representation of Euclidean space. In this new technique, the Euclidean space axes are transformed into a set of parallel axes arrangements where each variable is represented by one parallel axis. The points are then plotted as series of lines passing through the parallel axes. Thus, any deviation in terms of these lines from the normal density area will be then identified indicating fault situation. The only problem that matter is that, the technique does not correspond to any form of variable association. Besides, the method is also merely applicable for off-line application only. Therefore, even though the idea of MDV is quite impressive, but the current technical limitation does not provide sufficient credit for promoting the method as the main alternative solution in combating the issue considered in this study.

Although it is generally understood that each of the methods has its own strength and capability, nevertheless, the limitation that relates to the problem of the context proposed in this study forced this work to find another option. As a result, multidimensional scaling (MDS) has been found to be relevant and possess the potential elements to be explored further in the light of the scopes specified (discussed in details in Section 2.4).

2.3.4 Issue 4: Fault Diagnosis

Fault diagnosis provides the sense of urgency in finding the true underlying problem that pertains to the signal detected. As explained earlier, this may involve two main stages, firstly, identifying the potential faulty variables through contribution plots, and later, specifically diagnosing those variable candidates in truly specifying the correct faulty variable(s). The identification procedures have been discussed as previously, therefore, this sub-section focuses on the review on the diagnosing techniques only. Nonetheless, the scope of discussion is narrowed down directly on the essential elements that relate between fault detection and diagnosis methods.

From the literature, fault diagnosis methods can be broadly categorized into two main groups – quantitative and qualitative methods. Three main mechanisms reviewed under the quantitative-based tool are decomposition technique, pattern recognition and single channel event (SCE) index. Mason and Young, (2002) proposed using the decomposition technique, that divides the principal components into its individual original entities. However, these entities, in terms of variables, subject to be restricted by one another, thus the terms conditional and unconditional were introduced. On the other hand, pattern recognition focuses on differentiating the outputs into several specified classes that are characterized by the input conditions. Such works have been developed by Akbaryan and Bishnoi, (2001) specifically focusing on analyzing noisy input patterns for fault diagnosis. Devillez et al., (2004) introduced integrated fuzzy logic and classification system (a kind of pattern recognition techniques) that deals in finding the appropriate current states that should reflect the identified functioning operating modes. Singhal and Seborg, (2006) have analyzed a number of similarity-based parameters for matching measure application and applied to the Tennessee Eastman (TE) challenge process, where several successful classification performances were also developed. In a rather similar approach to this (in terms of objective), Yoon and MacGregor, (2001) utilized fault signature method for fault diagnosis. Their approach constitutes of quantitative comparison in terms of movement angle between the current and known fault vectors that obtained from the fault library. Unlike the previous methods which are all vector driven, SCE employs the values of T^2 and SPE statistics in calculating the diagnostic index (Ramaker et al., 2006). Another major difference is that this approach is solely applicable for sensor failure management only and each variable possesses their unique set of indices.

All of these techniques (except pattern recognition) show that their individual approach takes into consideration the previous information on fault detection, either in terms of vector or statistics measure, in delivering the diagnosis solution. It is therefore, the new monitoring system proposed in this study should also, in a similar manner, provide these basic elements, however, not necessarily in the same technical formulation and methodology. More importantly, the new multivariate model that suggested must somehow designed to be flexible and transformable that relates between the monitoring statistic or vector values (in the latent space) and the original variables (in the higher order dimension domain) for ease of signal interpretation and fault isolation.

Qualitative methods, however, specifically use knowledge-based system (KBS) and have introduced another extra advantage rather than diagnosis solution alone. In particular, the technique incorporates intelligent supervision mechanism which can execute fault detection, identification and diagnosis simultaneously and automatically with or without human intervention. Examples of such methods are given in the works of Norvilas et al., (2000) for continuous systems and Undey et al., (2003. 2004) for batch processes. The method was also proven functioning effectively for industrial monitoring (Skoglund et al., 2005; Chew et al., 2007). In particular, a KBS basically consists of process specific knowledge (quantitative and qualitative) and inference engine. Thus, unlike the idea of quantitative diagnostic tool, KBS directly accesses to the current process information in order to validate the abnormal event diagnosis. In other words, MSPM provides the basic detection scheme, whereas KBS supplies the related information, in terms of isolating fault causes and supervising corrective actions pertaining to the detected situation, through the real time searching and heuristic algorithms. Hence, the connection between MSPM and KBS can be perceived as loose or passive in relative to the quantitative techniques.

2.3.5 Issue 5: Data Formulation

Firstly, data formulation involves expanding the scopes of the data used in monitoring. Normally, off-line process history data of normal operating measurements are utilised to develop the multivariate model. There are also other forms of multivariate data model have been proposed for monitoring, for examples, state space model (discussed previously), external information in terms of batch-specific and run-specific data (Ramaker et al., 2002) as well as subspace identification (Treasure et al., 2004) in studying specifically the dynamical process behaviour. While in another instance, the information extracted from multivariate sensor (MacGregor et al., 2005) focuses on investigating the implications of monitoring and control impact by way of integrating molecular properties information and multivariate process data. Reis and Saraiva, (2006), on the other hand, utilizes heteroscedastic latent variable (HLV) in handling missing as well as uncertainty data. Lastly, the data-driven quality improvement (DDQI) technique proposed by Kano et al., (2008) combines optimized operating condition and quality measurement knowledge for monitoring. As a whole, this particular direction accentuates on the preliminary step before applying the multivariate technique.

The challenges forwarded by all these works are truly relevant. Nonetheless, the emphasis is limited according to the scopes of this study (corresponding to only the first and second steps of Figure 2.1) and they do not significantly change the original interpretation on the typical variable correlation measures (even though this study recognizes the contribution they have made in improving the monitoring performance). Thus, the main issue addressed in this study will be still unsolved as a result of adopting all of these techniques. Perhaps, these methods are more suitable to be considered in the next step of monitoring development level.

2.3.6 Issue 6: Automation System

The implementation of automatic monitoring system, on the other hand, comes at the last stage, where all the necessary monitoring elements have been well established (detection, identification, diagnosis and control are proven working) and ready for the real time industrial applications. In the case of PCA as an example, Kourti et al., (1996) have discussed some of the basic benefits and issues regarding of their real industrial experiences. In a rather extensive review, Miletic et al., (2004) have shared their knowledge on building an on-line monitoring system in the real time environment, where the technical and non-technical issues were raised and addressed.

In addition to this, Venkatasubramaniam et al., (2003a; 2003b; 2003c) have conducted a comprehensive study in comparing between various kinds of monitoring systems that are available including quantitative, qualitative and process history-based techniques. Their findings lead them to conclude that every method has their own individual and unique strengths and also weaknesses. Therefore, they proposed a hybrid system in order to overcome the limitations, and at the same time, reinforce the strength of every individual method. In other words, the hybrid-based system allows all the techniques to complement each other in terms of their individual advantages. In conjunction, they also suggested that this newly integrated system should also be supported by an intelligent supervisory system that suitable for large scale industrial applications.

2.4 MDS as An Alternative Solution for MSPM System

A brief description of MDS is presented in the first chapter covering its functionality, potentials and challenges with regard to the specified problem. This section, on the other hand, further discusses on the rationality of using MDS for MSPM system improvement that support the arguments made previously. The discussion, firstly presents the general concept of MDS as well as the specific mathematical background of classical multidimensional scaling (CMDS) as a core technique in the MDS group. Then, the reviews on using the MDS techniques in different kinds of monitoring systems are described. Next, a brief explanation on the interpretation and challenges of applying MDS in the light of JPMC is provided. Finally, a summary is given highlighting the justification on 'closing the gap' that specifically connected to the issue addressed in Chapter 1.

2.4.1 Conceptual Background of MDS

MDS has served various purposes and one of those relates to the exploratory data technique. The term exploratory basically connects to the analysis on the general structure of that particular multivariate data under investigation which later contributes to the establishment of conceptual perception and understanding. As far as this matter is concerned, the task involves reducing the high dimensional multivariate data into a lower dimensional model configuration that is much simpler for visual interpretation and sensible for further evaluation. From the literature survey, most of the early applications of MDS are originated from the field of social science, for instance, Torgerson, (1967) has briefly stated that:

"Given a set of stimuli, which vary with respect to an unknown number of dimensions, determine (a) the minimum dimensionality of the set, and (b) projections of the stimuli (scale values) on each of the dimensions involved."

The term 'stimuli' or stimulus are referred to those intangible entities that are significant to the non-technical research areas such as psychology and education. With regard to engineering applications, then, stimulus could be replaced by any measure of variables that could be directly computed using the appropriate devices wherever they are required. Kruskal and Wish, (1978), on the other hand, has defined MDS more towards to object-proximities-oriented basis:

"MDS, then, refers to a class of techniques. These techniques use proximities among any kind of objects as input........The chief output is a spatial representation, consisting of a geometric configuration of points, as on a map."

The word 'points' in that statement, indicate to the objects as the prime input, whereby this could be referred as those inanimate things such as observation numbers, materials or systems. In the later years, there was a trend, where, several authors have referred those 'proximities' as synonyms to 'distance' such as:

i. Chatfield & Collins, (1980)

"MDS is essentially concerned with finding a configuration of points (or individuals or objects) from information about the 'distances' between the points."

ii. Coxon, (1982)

"MDS refers to a family of models by means of which information contained in a set of data is represented by a set of points in a space. These points are arranged in such a way that geometrical relationships such as 'distance' between the points reflect the empirical relationships in the data."

iii. Cox, & Cox, (1994)

"A narrow definition of MDS is the search for a low dimensional space, usually Euclidean, in which points in the space represent the objects ....and such that the distances between the points in the space, match as well as possible the original dissimilarities."

iv. Borg, & Groenen, (1997)

"MDS is a method that represents measurements of similarity (or dissimilarity) among pairs of objects as distances between points of a low-dimensional multidimensional space".

v. Takane, (2003)

"MDS is a data analysis technique to locate a set of points in a multidimensional space in such a way that points corresponding to similar stimuli are located close together, while those corresponding to dissimilar stimuli are located far apart."

vi. Cox, (2005)

"MDS covers a variety of techniques, aimed at representing objects (or individuals) by a configuration of points in a space, usually Euclidean,..........., objects which are ‘similar’, in some sense, are expected to have their respective points in the space close to each other, while objects which are dissimilar would have their points further apart.”

In analysing those definitions, three important elements have been identified with regard to MDS – dissimilarity scales (proximities), spatial configuration and measure of fitness. With respect to dissimilarity measure, this particular factor is typically used as the correlation concept, whereby the points in a reduced dimensional space are arranged such that their distances correspond to the distances among the original data points in the original data space. In other words, the dissimilarity between two objects is a measure of ‘how dissimilar’ they are upon each other. The final form of dissimilarity measures is always written in a matrix structure as shown in the following:

$$\mathbf{\Delta}_{m\times m}=\begin{bmatrix} \delta_{1,1} & \delta_{1,2} & \cdots & \delta_{1,m} \\ \delta_{2,1} & \delta_{2,2} & \cdots & \delta_{2,m} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \delta_{m,1} & \delta_{m,2} & \cdots & \delta_{m,m} \end{bmatrix} \tag{2.38}$$

Where, $\mathbf{\Delta}_{m\times m}$= dissimilarity matrix with size m by m.

$\delta_{m,m}$= element of dissimilarity matrix at row m and column m.

Apart from dissimilarity measure, another important element, which is essential in developing any set of MDS score, is spatial configuration. According to Torgerson, (1967), any of geometric spaces could potentially be used as the basic spatial model for a MDS procedure. However, he also claimed that the Euclidean space in terms of Cartesian coordinates is possibly the only one which has been extensively utilised in this regard. This is perhaps, it possess several advantages as listed as follows:

i. It is the most familiar.
ii. Graphical representation is convenient.
iii. It has a strong theoretical and conceptual simplicity.
iv. It has direct and simple mathematical properties.

A Cartesian coordinate system is a set of pair wise perpendicular straight lines (coordinate axes). Each axis has the same units and all axes intersect at one point, the origin, O. If $\mathbf{X}_E$ is a Cartesian $n_E \times m_E$ coordinate matrix, then, it actually denotes that:

i. a set of n_E points in m_E-dimensional space, or
ii. the coordinates of the n_E-points relative to m_E-Cartesian coordinate axes.

In addition, Green and Carroll, (1976) have stated that a Euclidean space of m_E-dimensions is the collection of all n_E-component vectors for which the operations of vector addition and multiplication by a scalar are permissible. Moreover, for any two vectors in the space, there is a non-negative number, called the Euclidean distance between the two vectors. In particular, the Euclidean distance between two points, i and j in $\mathbf{X}_E$ is the length of a straight line connecting points i and j in $\mathbf{X}_E$. As far as the Euclidean space is concerned, certain properties must be obeyed during which, the transformation process is taken place as shown as follows (Coxon, 1982):

i. Non-negativity and equivalency
   a. $\delta_{ij} \geq 0$ for all points *i* and *j*.
   b. $\delta_{ij} = 0$ if and only if *i* coincides with *j*.
ii. Symmetry: $d_{ij} = d_{ji}$ for all points *i* and *j*.
iii. Triangle Inequality
   a. $d_{ij} = d_{ik} + d_{kj}$ for all points *i,j* and *k* (point *k* lies on line *ij*).
   b. $d_{ij} < d_{ik} + d_{kj}$ for all points *i,j* and *k* (point *k* lies off line *ij*).
   c. This clearly excludes the possibility that $d_{ij} > d_{ik} + d_{kj}$ for any case.

Lastly, MDS models also require a set of fitness measure when comparing the redeveloped and original data configuration. According to Torgerson, (1967), this is basically means that if the distance model yields absolute distances conceptually, it is also meaningful to ask whether the distances is really exist in a real space, which subject to the number of dimensional space that chosen.

Coxon, (1982), has divided the measure of fits into two separate forms (in both terms, $d_{E\,jk}$ and $d_{jk}$are the redeveloped and original dissimilarity measure respectively):

i. The difference $\left(d_{E_{jk}} - d_{jk}\right)^2$, which represents the basis of 'badness-of-fit', since the greater the discrepancy between a solution and the data, the greater will be the differences.

ii. The scalar product $\left(d_{E_{jk}} \hat{d}_{jk}\right)$, which forms the basis of 'goodness-of-fit', since the greater the covariance between data and the solution, the greater will be the scalar products, the better configuration is developed.

As far as those descriptions are concerned, the extensions of MDS concept can be summarized as follow:

i. Normally, points in the space represent the objects rather than stimulus or variables. Yet, it doesn't mean that the variables cannot be plotted alternatively.

ii. Dissimilarity scales are fundamental for data input, whereby, Euclidean distance usually being chosen for distance model. Yet, other measures of distances are also allowable to be analyzed in this respect, despite that, the Cartesian coordinate is still important to be applied.

iii. The multidimensional space is still critically to be the subject for optimization, whereby, a low-dimensional space is usually being the main objective.

2.4.2 Mathematical Fundamentals of MDS

Generally, there are two basic methods of MDS namely as classical multidimensional scaling (CMDS) and non-metric multidimensional scaling (NMDS) that are popularly utilised (Torgerson, 1967; Coxon, 1982; Green et al., 1989; Cox and Cox, 1994; Borg and Groenen, 1997; Cox, 2005). The first uses metric-based approach, whereas the second applies non-metric mechanism (as the name implied). This study adopts only the first method because it holds a strong relationship with PCA. In particular, the fundamental of this procedure is originated from the Young –Householder Theorem, (1938).

Thus, the following descriptions lay out the mathematical derivation on the transformation of the original multivariate data, **X**, that can be reproduced back into a new form of compressed multivariate model, $\mathbf{X_E}$, particularly using the detailed procedures of CMDS and that also explains the previous theorem. As the technique uses the metric-based approach, the original data is subject to be standardized and works on the ground of Euclidean space (Green et al., 1989). The standard CMDS procedures considered in this study is originally proposed by Cox and Cox, (1994), Borg and Groenen, (1997), and Cox, (2005), where the focus is on constructing the scores in terms of sample distribution. Nevertheless, it has been radically modified for developing the scores by means of variable configurations (which has been particularly emphasized in this study) as the following:

Step 1. Determination of a squared dissimilarity matrix, $\mathbf{\Delta^2}$ based on $\mathbf{X}_{m\times n}$ (m: variables, n: samples). The size of $\mathbf{\Delta^2}$ is 'm' by 'm'.

Step 2. Computation of a double-centred matrix, $\mathbf{B}_\Delta$, on $\mathbf{\Delta^2}$. This is important in order to relocate the origin of $\mathbf{\Delta^2}$ into the centre (Torgerson, 1967) so that a unique set of score configuration can be obtained eventually.

Step 3. Decompose $\mathbf{B}_\Delta$ into eigen basic structure by using singular value decomposition (SVD) method.

Step 4. Development of multivariate scores based on Cartesian coordinates, $\mathbf{X_E}$.

In Step 1, the squared dissimilarity matrix, $\mathbf{\Delta^2}$, is determined simply by using a particular distance measure. Two commonly used distance measures are Euclidean and City-block distances (Cox and Cox, 1994; Cox, 2005) as shown in the following:

Euclidean distance: $$\delta_{ij} = \{\sum_a (x_{ia} - x_{ja})^2\}^{\frac{1}{2}} \quad (2.39)$$

City-block distance: $$\delta_{ij} = \sum_a | x_{ia} - x_{ja} | \quad (2.40)$$

where δ_{ij} is the dissimilarity measure between objects i and j, x represents the original Cartesian coordinate of those objects (variables), and a is the number of axes in the original space.

$\mathbf{B}_\Delta$ is then obtained by using Equation (2.41).

$$\mathbf{B}_\Delta = -\frac{1}{2}\mathbf{J}_m \Delta^2 \mathbf{J}_m \quad \text{where} \quad \mathbf{J}_m = (\mathbf{I}_m - \mathbf{1}_m \mathbf{1}_m'/m) \quad (2.41)$$

where $\mathbf{I}_m$ is an identity matrix of size $m\times m$, while $\mathbf{1}_m$ is a vector of size m with all elements equal to 1. Additional remarks on $\mathbf{J}_m$ are provided in **Appendix A**.

In the following step, $\mathbf{B}_\Delta$ will be decomposed into $\mathbf{UVU^T}$ where $\mathbf{U}$ and $\mathbf{V}$ are, respectively, the orthonormal eigenvectors and a diagonal matrix containing the eigenvalues of $\mathbf{B}_\Delta$. Finally, the reconstructed multivariate scores can be calculated by using Equation (2.42).

$$\mathbf{X_E}=\mathbf{U_+V_+^{0.5}} \tag{2.42}$$

where $\mathbf{V_+^{0.5}}$ is a diagonal matrix with all positive elements of $v^{0.5}$ and $\mathbf{U}_+$ is the corresponding sets of eigenvectors.

The degree of relationship between the retained dimensions and the original data can be assessed based on Equation (2.43) (Cox and Cox, 1994).

$$\rho = \sum_{i=1}^{p} v_i \Big/ \sum_{i=1}^{m} v_i \tag{2.43}$$

In Equation (2.43), v_i is the ith eigenvalue of $\mathbf{B}_\Delta$ (arranged in descending order), p is the number of selected dimensions and m is the total number of variables. The higher the value of ρ, the stronger **X** is mapped or reproduced by **Y** in terms of its inter-distance scales. Figure 2.2 summarizes all the CMDS procedures discussed in this section with an example demonstrated by Borg and Groenen, (1997). In particular, the redevelopment of the scores was performed based on the first four items which regarded to facial expression analysis as listed as follow:

1.0 Item 1: Grief at death of mother (variable 1 – V1)
2.0 Item 2: Savouring a coke (variable 2 – V2)
3.0 Item 3: Very pleasant surprise (variable 3 – V3)
4.0 Item 4: Maternal love – baby in arms (variable 4 – V4)

According to them, 30 students (objects) were asked to rate the proximities of those four items initially before it was transferred and finalised into MDS distances form that used in Figure 2.2. The final of the new coordinates obtained from the example shown in Figure 2.2 are shown graphically in Figure 2.3.

Step 1: Defining the dissimilarity matrix.

$$\Delta = \begin{bmatrix} 0 & 4.05 & 8.25 & 5.57 \\ 4.05 & 0 & 2.54 & 2.69 \\ 8.25 & 2.54 & 0 & 2.11 \\ 5.57 & 2.69 & 2.11 & 0 \end{bmatrix} \text{, so that } \Delta^2 = \begin{bmatrix} 0 & 16.40 & 68.06 & 31.02 \\ 16.40 & 0 & 6.45 & 7.24 \\ 68.06 & 6.45 & 0 & 4.45 \\ 31.02 & 7.24 & 4.45 & 0 \end{bmatrix}$$

Step 2: Applying double centering.

$$\mathbf{B}_\Delta = -\frac{1}{2}\mathbf{J}_\mathbf{n}\boldsymbol{\Delta}^2\mathbf{J}_\mathbf{n}$$

$$= -\frac{1}{2}\begin{bmatrix} +\frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & +\frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & +\frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & +\frac{3}{4} \end{bmatrix} \begin{bmatrix} 0 & 16.40 & 68.06 & 31.02 \\ 16.40 & 0 & 6.45 & 7.24 \\ 68.06 & 6.45 & 0 & 4.45 \\ 31.02 & 7.24 & 4.45 & 0 \end{bmatrix} \begin{bmatrix} +\frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & +\frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & +\frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & +\frac{3}{4} \end{bmatrix}$$

$$= \begin{bmatrix} +20.52 & +1.64 & -18.08 & -4.09 \\ +1.64 & -0.83 & +2.05 & -2.87 \\ -18.08 & +2.05 & +11.39 & +4.63 \\ -4.09 & -2.87 & +4.63 & +2.33 \end{bmatrix}$$

Step 3: Computing eigen decomposition of $\mathbf{B}_\Delta$.

$$\mathbf{B}_\Delta = \mathbf{U}\mathbf{V}\mathbf{U}^T$$

$$\begin{bmatrix} +0.77 & +0.04 & +0.50 & -0.39 \\ +0.01 & -0.61 & +0.50 & +0.61 \\ -0.61 & -0.19 & +0.50 & -0.59 \\ -0.18 & +0.76 & +0.50 & +0.37 \end{bmatrix} \begin{bmatrix} 35.71 & 0.00 & 0.00 & 0.00 \\ 0.00 & 3.27 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & -5.57 \end{bmatrix} \begin{bmatrix} +0.77 & +0.01 & -0.61 & -0.18 \\ +0.04 & -0.61 & -0.19 & +0.76 \\ +0.50 & +0.50 & +0.50 & +0.50 \\ -0.39 & +0.61 & -0.59 & +0.37 \end{bmatrix}$$

Step 4: Calculating the Cartesian coordinate matrix, $\mathbf{X}_E$.

$$\mathbf{X}_E = \mathbf{U}_+\mathbf{D}_+^{0.5} \;\Rightarrow\; \begin{bmatrix} +0.77 & +0.04 \\ +0.01 & -0.61 \\ -0.61 & -0.91 \\ -0.18 & +0.76 \end{bmatrix} \begin{bmatrix} 35.71 & 0 \\ 0 & 3.27 \end{bmatrix}^{0.5} = \begin{bmatrix} +4.60 & +0.07 \\ +0.06 & -1.10 \\ -3.65 & -0.34 \\ -1.08 & +1.37 \end{bmatrix}$$

Figure 2.2: A numerical example for CMDS

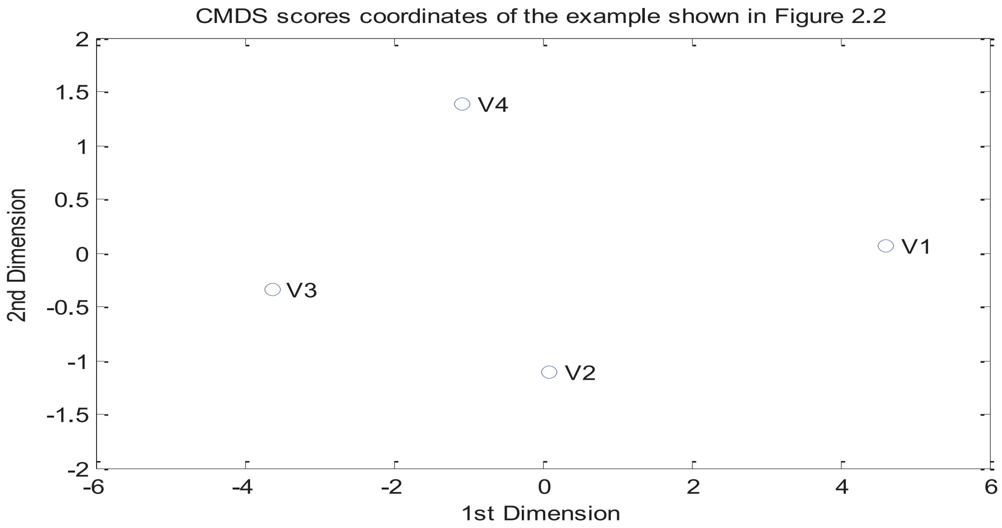

Figure 2.3: The reproduction of CMDS scores that demonstrated in Figure 2.2

From Figure 2.3, it can be seen that the new score coordinates reflect the original dissimilarity scales, Δ (as shown in Table 2.4), that was used to build the scores.

Table 2.4: Dissimilarity matrix based on Euclidean scales used in Figure 2.2

	V1	V2	V3	V4
V1	0	4.05	8.25	5.57
V2	4.05	0	2.54	2.69
V3	8.25	2.54	0	2.11
V4	5.57	2.69	2.11	0

In particular, from Table 2.4, it was indicated that V3 has the largest distance from V1 (8.25) in relative to the other two points (V2 and V4). As a result, the reproduction of the scores in Figure 2.3 also corresponds to these situations, where the location of V3 will be configured far away from V1 in contra to V2 and V4. The other coordinates also follow the similar order particularly based on the original dissimilarity scales that applied. In reflecting to the interpretation of those items, it is naturally accepted and well understood that facial expression for someone who is grieving at death of mother (V1) will radically different from the facial look relating to one that enjoying a very pleasant surprise moment (V3). Meanwhile, those of facial expressions pertaining to savouring a coke (V2), enjoying a pleasant surprise scene (V3) and also maternal love (V4) should be easily rationalise as demonstrating close distances with each other, as they are all representing positive emotional attitude.

CMDS typically utilises 'strain' parameter in specifically measuring the degree of mapping between the new and original dissimilarity scales. Torgerson, (1967), explained that the term is strongly associated with scalar product moment and conceptually defined it as the comparison between the total variance in the original scalar product matrix, $\mathbf{B}_{\Delta}$, with the total variance of a derived scalar product matrix from the new configuration of Euclidean space coordinate, $\mathbf{B}_E$, as depicted in Equation (2.44).

$$I = \sum(\mathbf{B}_{\Delta} - \mathbf{B}_{\mathbf{E}})^2 \quad (2.44)$$

Fundamentally, the scalar product is defined as:

$$\mathbf{a'b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

$$= \sum_{k=1}^{n} a_k b_k \quad (2.45)$$

whereby, $\mathbf{a} = [a_1\ a_2\ \dots\ a_n]^T$ and $\mathbf{b} = [b_1\ b_2\ \dots\ b_n]^T$

The scalar product also represents the magnitude or the length of a vector in Cartesian coordinates, in which, it may also be written as:

$$\|\mathbf{a}'\| = (\mathbf{a'a})^2 \quad (2.46)$$

As much as the Euclidean distance is concerned, there are two ways, where, the relationship of the scalar product matrix of the Cartesian coordinates with the Euclidean distance is evident.

Torgerson, (1967), Coxon, (1982) as well as Borg and Groenen, (1997) have used the cosine law to establish this association, in which the procedures are briefly shown as follows:

i. Let d_{ij}, d_{ik} and d_{jk}be the Euclidean distances between *'n'* points.
ii. Then, $\mathbf{B}_i$ is a symmetric matrix with elements

$$b_{jk} = \frac{1}{2}\left(d_{ij}^2 + d_{ik}^2 - d_{jk}^2\right) \quad (2.47)$$

The element b_{jk} can be interpreted as the scalar product of the vectors from point 'i'' to point 'j' and 'k'.

iii. From the law of cosine theorem,

$$d_{jk}^2 = d_{ij}^2 + d_{ik}^2 - 2d_{ij}d_{ik}\cos\theta_{ijk} \tag{2.48}$$

and after some rearrangements, eventually, the scalar product is directly obtained by

$$d_{ij}d_{ik}\cos\theta_{ijk} = \frac{1}{2}\left(d_{ij}^2 + d_{ik}^2 - d_{jk}^2\right) \tag{2.49}$$

iv. From this derivation, it obviously shows that

$$b_{jk} = d_{ij}d_{ik}\cos\theta_{ijk} \tag{2.50}$$

2.4.3 Connections between PCA and MDS

A search was also executed for investigating the relationship between PCA and MDS. This relationship is viewed from the close fundamental algorithms between conventional PCA and CMDS procedures. Several authors have illustrated different ways in describing the relationship between PCA and MDS using algorithm manipulations approach. Cox and Cox, (1994) started the procedure by defining the scalar product matrix, **B** (or specifically the dissimilarity matrix of δ_{ij}^2), $\mathbf{B} = \mathbf{X}\mathbf{X}^T$, where **X** is an $n \times m$, data matrix assumed to be mean centred. By applying the SD operation on **B**, the following are obtained:

$$\mathbf{B}\mathbf{u}_i = \lambda_i \mathbf{u}_i \tag{2.51}$$

$$\mathbf{X}\mathbf{X}^T\mathbf{u}_i = \lambda_i \mathbf{u}_i \tag{2.52}$$

$$\mathbf{X}^T\mathbf{X}\mathbf{X}^T\mathbf{u}_i = \lambda_i \mathbf{X}^T\mathbf{u}_i \tag{2.53}$$

$$\mathbf{C}\mathbf{q}_i^* = \lambda_i \mathbf{q}_i^* \tag{2.54}$$

where,

$$\mathbf{C} = \mathbf{X}^T\mathbf{X} \tag{2.55}$$

$$\mathbf{q}_i^* = \mathbf{X}^T\mathbf{u}_i \tag{2.56}$$

In this particular derivation, equation (2.56) is known generally as the typical SD structure in the eigen structure procedure, where **C** represents the minor product moment of **X** and $\mathbf{q}_i^*$ stands for the eigenvectors of that minor product moment. However, $\mathbf{q}_i^*$ should be normalized by $\lambda_i^{-0.5}$, as to get the true loading factors as in PCA (orthogonal), which are derived as follows:

Normalizing $\mathbf{q}_i^*$,

$$\mathbf{q}_i = \mathbf{X}^T\mathbf{u}_i\lambda_i^{-0.5} \tag{2.57}$$

Therefore,

$$\mathbf{q}_i^T\mathbf{q}_i = \left[\mathbf{X}^T\mathbf{u}_i\lambda_i^{-0.5}\right]^T\left[\mathbf{X}^T\mathbf{u}_i\lambda_i^{-0.5}\right] \tag{2.58}$$

$$= \lambda_i^{-0.5}\lambda_i^{-0.5}\mathbf{u}_i^T\mathbf{X}\mathbf{X}^T\mathbf{u}_i \tag{2.59}$$

$$= \lambda_i^{-1}\lambda_i^{+1} = 1 \tag{2.60}$$

Besides, Jackson, (1991) has also shown that major product moment, $\mathbf{XX}^T$, and minor product moment, $\mathbf{X}^T\mathbf{X}$, of $\mathbf{X}$ share the same eigenvalues. In his method of explanation, the major product moment is called as Q-analysis, whereas the other one is called as R-analysis. He also stated that Q-analysis could be used to construct the scores for variables and, on the other hand, R-analysis is useful normally for developing the normal observations scores. Those steps are shown in the following procedures:

$$\text{SVD of } \mathbf{X}^T\mathbf{X} = \mathbf{QDQ}^T \tag{2.61}$$

$$\text{SVD of } \mathbf{XX}^T = \mathbf{UDU}^T \tag{2.62}$$

Then, the PC scores for the observations and variables are given individually by:

$$\mathbf{Z}_O = \mathbf{XQD}^{-0.5} \tag{2.63}$$

$$\mathbf{Z}_V = \mathbf{X}^T\mathbf{UD}^{-0.5} \tag{2.64}$$

In analysing this procedure, it can be noted that PCA is capable of plotting scores either for observations or variables by way of algorithm scaling. Even though, the usual PC scores are normally computed by:

$$\mathbf{Z} = \mathbf{XQ} \tag{2.65}$$

In using equation (2.65), it can be shown that PCA has its direct connection with CMDS. At the beginning, after being performed under SD structure, then, the first PC score is given by:

$$\mathbf{z}_{01} = \mathbf{Xq_1} \tag{2.66}$$

By inserting equation (2.57) into equation (2.66):

$$\mathbf{z}_{01} = \mathbf{XX}^T\mathbf{u_1}D_1^{-0.5};\ \mathbf{q_1} = \mathbf{X}^T\mathbf{u_1}\ D_1^{-0.5}$$

$$\mathbf{z}_{01} = D_1\mathbf{u_1}D_1^{-0.5};\ \ \mathbf{XX}^T\mathbf{u_1} = D_1\mathbf{u_1}$$

$$\mathbf{z}_{01} = \mathbf{u_1}D_1^{0.5} \tag{2.67}$$

which validates that equation (2.67) is the typical matrix multiplication to obtain the new score coordinates as in CMDS. The presented procedures are valid only for the first PC scores, therefore, the same procedures should be performed for other dimension scores. This shows that, by initiating some modifications on the MDS algorithm, the results will lead to the similar outcomes as in PCA.

In addition, another way to find the PC scores is by using the SVD on $\mathbf{X}$ instead of $\mathbf{X}^T\mathbf{X}$, as given by the following equations:

$$\text{SVD of } \mathbf{X} = \mathbf{P}\mathbf{D}^{0.5}\mathbf{T}^T; \tag{2.68}$$

where,
$$\mathbf{T} = \mathbf{Q} \tag{2.69}$$

Therefore, the usual PC scores can be computed by:

$$\mathbf{Z} = \mathbf{X}\mathbf{T} = \mathbf{X}\mathbf{Q} \tag{2.70}$$

In particular, equation (2.68) could also be related back to the minor product moment of $\mathbf{X}$, which is shown as follows:

Minor product moment;
$$\mathbf{C} = \mathbf{X}^T\mathbf{X} \tag{2.71}$$

Therefore,
$$= \left[\mathbf{P}\mathbf{D}^{0.5}\mathbf{Q}^T\right]^T\left[\mathbf{P}\mathbf{D}^{0.5}\mathbf{Q}^T\right] \tag{2.72}$$

$$= \mathbf{Q}\mathbf{D}^{0.5}\mathbf{P}^T\mathbf{P}\mathbf{D}^{0.5}\mathbf{Q}^T \tag{2.73}$$

$$= \mathbf{Q}\mathbf{D}\mathbf{Q}^T \tag{2.74}$$

2.4.4 Previous Works on MDS-based Monitoring Systems

The original work of using MDS for process monitoring was presented by Cox, (2001) and also described in Cox, (2003). In both of the works, even though MDS was introduced rather superficially but the idea was very inspiring because it spurs a new perspective in understanding the nature on the variable correlations. In his works, two different methods, by way of score configurations, were proposed. In particular, the first summarized the multivariate data by means of observation-sample profiles just exactly similar to the PCA outputs. Thus, whenever a fault situation occurred, the corresponding abnormal samples will either deviate in a great magnitude or simply gradually from the normal cluster (based on the dimensional scatter plot). In another approach, however, the studies have introduced the variable scores instead, specifically in providing the status of monitoring progression. In this particular approach, the variables that correspond to the particular malfunction condition are also responding in the same way as the previous method, but with one unique advantage.

Unlike to the former method, the movements made by those abnormal variable scores are not just restricted in showing the magnitude of deviation, but more importantly it is also actually projecting the information on violation of the normal variables' coordination (in terms of multivariate scores). Thus, reproducing the scores by way of variables is more informative rather than the conventional score configurations (observation samples) from process monitoring perspective. Thus, this shows that the variable-based form of scores can also be potentially utilized for monitoring application. However, this study did not propose any kind of monitoring statistics that can be compared against the standard PCA-based monitoring performance.

In another study, Matheus et al., (2006) has applied multiple linear regression (MLR) in projecting the MDS scores for on-line monitoring system. They used the topographic mapping and clustering operation as the basis for fault detection mechanism. As similar to the previous work, this study also did not utilize any monitoring statistics as the basic measure for fault detection. Thus, the work is lacking in validity because it cannot be used for comparison with the conventional MSPM performance. Besides, there was no specific explanation on the assumption that they have used in developing the dissimilarity measures, particularly in relation to the JPMC perspective.

This review also has considered other techniques that could be perceived as the main competitors for MDS (beside of PCA). Firstly, Kano et al., (2002) proposed using the dissimilarity index (DISSIM method) for monitoring changes in the operating condition based on sample distribution profile. A fault is detected when the proposed monitoring index violates its control limit setting. Nonetheless, in comparison to MDS as well as PCA, this technique is not actually a tool for data compression (Kano et al., 2002), where the index is calculated directly from the input data. Therefore, once a fault event is detected, the interrogation has to deal with the complexity of high dimensional data. Lee et al., (2004), on the other hand, introduced hierarchical clustering strategy to differentiate various operation modes through *k*-means clustering. Even though the idea can be connected to the concept of T^2, nevertheless, the approach cannot be used in extracting the information on the variable correlations (SPE). Therefore, this survey finds that MDS is the closets and comparable technique to the PCA method, where the justifications are given in the next subsection.

2.4.5 Justification of Applying MDS in the MSPM Framework

This sub-section basically explains on how the MDS approach can be fitted into the elements set by JPMC. First of all, MDS has the ability to reproduce the scores either in terms of observation sample or variable configurations as proved by Cox, (2001) and Cox, (2003). The former has used the dissimilarity scales by way of sample observation correlations whereas the latter utilizes the dissimilarity measures by means of variable relationships. In corresponding to the third element of JPMC, which suggests that any monitoring system should be executed and reflected from the ground of variable correlations, thus, the second is perhaps the most suitable technique that should be chosen in this respect. Compared to the fundamental of conventional system, MDS, particularly by using the second approach, uses dissimilarity scales instead of variance-covariance basis. Enforcing of using the first approach, however, creates two undesirable situations. Firstly, the dissimilarity scales in terms of observation sample structure cannot be directly used in analyzing the variable correlations (as typically being developed by the previous works on using MDS for monitoring). Even so, another transformation function should be used and validated in establishing the relationship between the observation dissimilarity scales and variable dissimilarity measures, thus this introduces more complexity (sometimes conflicts) into the algorithms.

The second reason is that the variable scores can be used directly in obtaining the monitoring statistics which are similar in concept to T^2 and SPE. In particular, T^2 basically provides the magnitude of deviation of the scores with respect to the mean (in the reduced dimensional PC space). In the case of MDS, however, measuring the distance of the deviated variable scores from the origin of the score plots or from the other means (such as individual points of NOC scores) may also deliver the same kind of information with regard to T^2. Besides, one of the main advantages of variable-score-based approach is that the changes in the score structure also depict changes in the variable correlations. Therefore, measuring the consistency of the dissimilarity scales between the current and NOC configurations can also be regarded as providing the similar information provided by the SPE objective. Hence, this suggestion addresses the first element of JPMC rules.

The main problem of applying the variable-score-based approach is that it cannot directly construct the distribution profile (because the scores themselves are indicating points of variables and not on sample distributions). Besides, the scores cannot be projected directly based on the variable vectors because the original MDS algorithms do not provide any kind of loading factor as normally utilised by PCA. Moreover, the creation of dissimilarity scales is very rigid as it necessitates the monitored data to be embedded in the form of a matrix with the same size as the original NOC data. Thus, all these drawbacks seem quite problematic especially when dealing with the on-line measurement data. In order to overcome all of these limitations, the moving window (MW) mechanism suggested by Kano et al. (2000; 2001; 2002) is proposed as the solution. The advantage of this approach lies in its ability to constantly construct the monitored data matrix of the same size as the original NOC data by using a moving window of the process operation data. Consequently, the relationships between the current and the previous data can also be utilized for dynamic monitoring (even though this is not the main issue concerned in this study). As a result, after analyzing a few samples, a set of moving-window-based samples that originated from the variable-score-based configurations can be established and perhaps may eventually generate the distribution profile that required for setting the control limits specification. Thus, this can be later employed as the basis in responding to the second element suggested by JPMC.

The final element of JPMC states that the system should be able to conduct fault identification and diagnosis following the detected fault event. As discussed earlier, the proposed scheme of monitoring progression based on the MDS approach should be performed by means of quantitative comparison between the current variable scores location and the original NOC scores configuration. This comparison can be objectively viewed as an 'error' measure, thus the concept of SPE formulation (or simply 'squared errors') is relevant in this context. As far as this matter is concerned, the contribution plot technique via SPE as shown in Equation 2.21 can be applied at the same time. Besides, MDS is also able to produce the scores which constitute of vectors and also facilitate for monitoring statistics calculation (as explained earlier). Thus, this method can also be potentially integrated with other quantitative or qualitative diagnosis tools for the purpose of diagnosis application.

2.5 Summary

It this chapter, the background of the conventional as well as the recent developments of MSPM are critically reviewed. Besides, some earlier studies on using MDS for monitoring are also discussed. Justifications are then forwarded to support the proposition on embedding the MDS technique into the MSPM frameworks. More specifically, the variable-score-based approach is selected as the main methodology that is also consistent with the elements required by JPMC.

CHAPTER 3

CASE STUDIES AND PCA BASED MONITORING PERFORMANCES

3.1 Introduction

A simulated Continuous Stirred Tank Reactor with Recycle (CSTRwR) system is used to demonstrate the proposed MDS based monitoring systems and compare with the conventional PCA based monitoring performances. The system has been used by Zhang (2006) to demonstrate the improved fault diagnosis technique by using multiple neural networks. The simulation was developed in Zhang, (1991) under normal and various faulty operating conditions based on a mechanistic model in the form of differential and algebraic equations of material balance, energy balance, and reaction kinetics (as presented in **Appendix B**). Some of the equations are in nonlinear form due to nonlinearities in reaction kinetics and valves. In the subsequent discussions, Section 3.2 describes the CSTRwR process including the NOC and various sets of fault cases, while Section 3.3 summarise the whole chapter.

3.2 CSTRwR System

3.2.1 Process Descriptions

The schematic diagram of a simulated CSTRwR is shown in Figure 3.1 (Zhang, 2006). This system conducts an irreversible heterogeneous catalytic exothermic reaction between reactant A and product B. The process is installed with three separate control loops, which consists of tank temperature, tank level and recycling flow variables, in order to maintain the product concentration. In particular, the cold water flow is adjusted through a cascade system corresponding to the changes in the reactor temperature. The reactor level, on the other hand, is maintained by controlling the flow rate of product. Lastly, the product composition in the reactor is indirectly controlled by manipulating the recycle flow rate. There are ten on-line measured process variables and three controller outputs have been identified for monitoring as listed in Table 3.1 (Zhang, 2006).

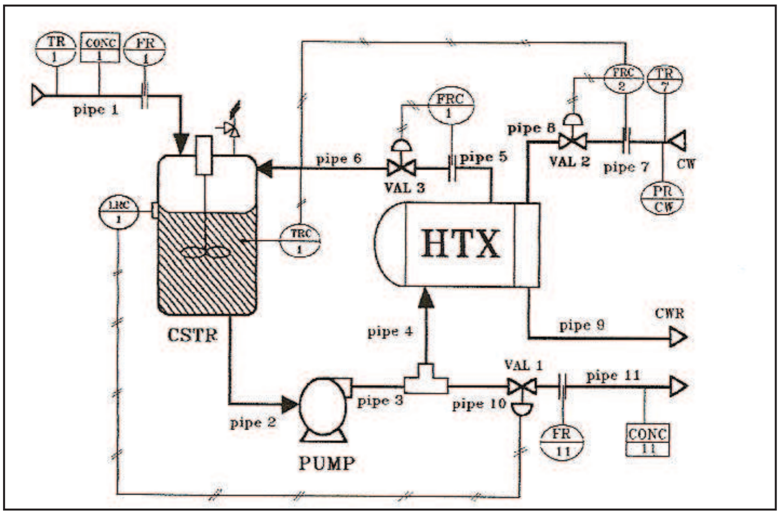

Figure 3.1: CSTRwR system

Table 3.1: List of variables in the CSTRwR system for monitoring

Process			Instruments		
No.	Variables	Variable Names	No.	Variables	Variable Names
1	V1	Tank temperature	11	V11	Controller 1
2	V2	Tank level	12	V12	Controller 3
3	V3	Feed temperature	13	V13	Controller 2
4	V4	Inlet flow rate			
5	V5	Recycle flow rate			
6	V6	Outlet flow rate			
7	V7	Cooling water flow rate		-	
8	V8	Product concentration			
9	V9	Feed concentration			
10	V10	Heat exchanger entrance pressure			

3.2.2 NOC Samples and PCA Monitoring Performances

In the first phase of monitoring, a set of NOC data containing 100 samples of steady state data was obtained from simulation (the profiles are provided in **Appendix C**). Before applying the PCA procedures, each of the variables in the NOC data was standardized to zero mean and unit variance. Figure 3.2 shows the accumulated data variances explained against the number of principal components (PCs). The result shows that almost 70%, 80% and 90% of the total variances are represented by using 3, 5 and 7 PCs respectively. In other words, the higher number of PCs being selected, the more data variances are captured by the PCA model. This indirectly indicates that the process tends to be non-linear as a relatively large number of PCs are required to represent sufficiently high data variations. PCA models with 3, 5 and 7 PCs are developed based on the particular set of NOC data.

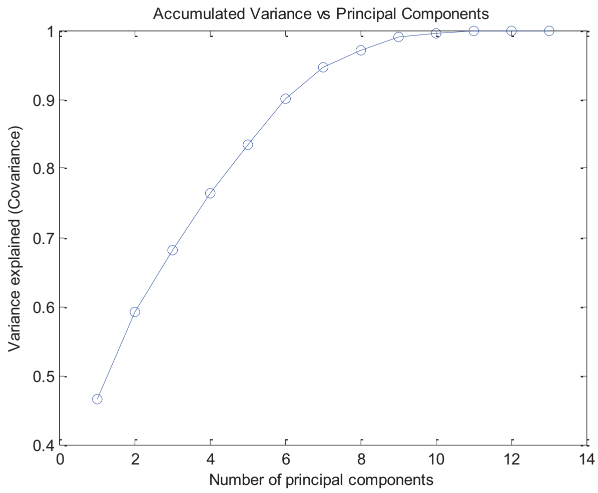

Figure 3.2: Accumulated data variations explained by different PCs for the CSTRwR

Then, monitoring limits were calculated for the T^2 and SPE statistics for both the PCA models. Figure 3.3 shows the T^2 and SPE statistics from the PCA model with 3, 5 and 7 PCs on the NOC data. The results show that all the T^2 monitoring statistics are below the monitoring limits, whereas only small amount of SPE statistics are slightly higher than the 99% control limit. This means that the three PCA models are truly represent the normal operation behaviour.

In order to evaluate the robustness of the monitoring limits, another set of NOC data (the second set of NOC) containing of 50 samples were also collected. The monitoring statistics are depicted in Figures 3.4 respectively for the 3, 5 and 7 PCs models. The results from Figure 3.4 show that all the monitoring statistic progressions are also below the control limits. Thus the developed PCA models successfully classify the new NOC data as being normal. These findings indicate that the monitoring systems do not give false alarms when applied to unseen NOC data. It was also observed that there was no significant difference between the trends in both of the PCA models, as far as NOC data is concerned.

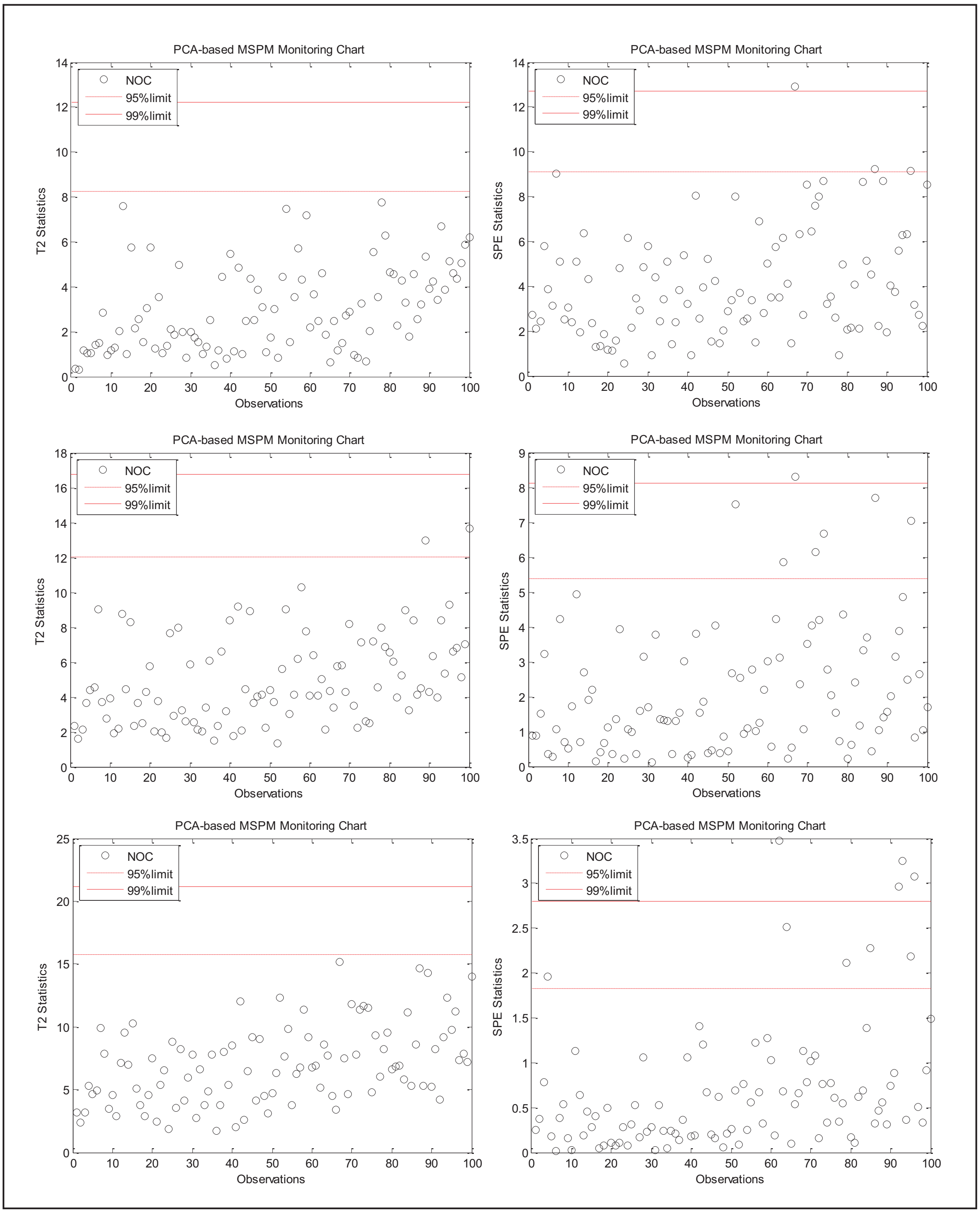

Figure 3.3: Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) on the original NOC data

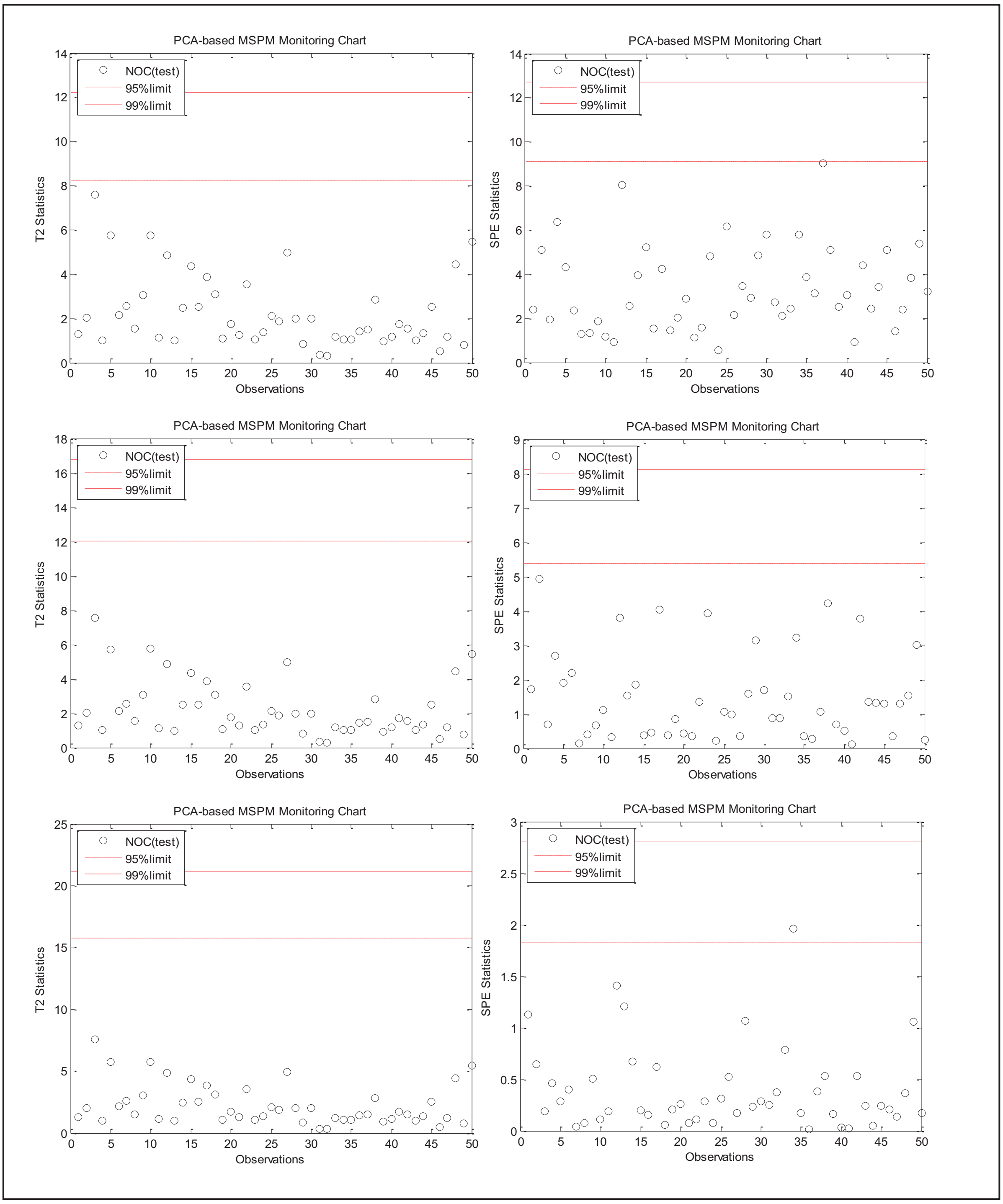

Figure 3.4: Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) on the testing NOC data

3.2.3 Fault Cases and PCA Monitoring Performances

The system also subjects to be affected from several malfunction conditions as summarized in Table 3.2.

Table 3.2: List of abnormal operations in CSTRwR

Fault Cases	Fault Causes
1	Pipe 1 blockage
2	External feed-reactant flow rate too high
3	Pipe 2 or 3 is blocked or pump fails
4	Pipe 10 or 11 is blocked or control valve 1 fails low
5	External feed-reactant temperature abnormal
6	Control valve 2 fails high
7	Pipe 7, 8, or 9 is blocked or control valve 2 fails low
8	Control valve 1 fails high
9	Pipe 4, 5, or 6 is blocked or control valve 3 fails low
10	Control valve 3 fails high
11	External feed-reactant concentration too low

For each fault presented in Table 3.2, both abrupt and incipient faults are considered. An abrupt fault indicates a sudden change (or step change) in a process variable or parameter and typically it maintains over the operation time until the cause is completely removed. Detecting this kind of malfunctions should be easy for any multivariate monitoring system as the deviations are usually very obvious. On the other hand, an incipient fault depicts a kind of fault that gradually deviates from the normal setting. Thus, the monitoring system typically takes a while to detect these particular abnormal behaviours. In particular, all the faults were introduced at sample 2 and the sampling time was fixed at 4 seconds. As examples, six cases, which are abrupt fault 6 (F6a), incipient fault 6 (F6i), abrupt fault 9 (F9a), incipient fault 9 (F9i), abrupt fault 11 (F11a) and incipient fault 11 (F11i), are selected to demonstrate the monitoring performance using the PCA-based systems.

In the first kind of abnormal operations, which corresponded to both F6a and F6i, the root of the problem is coming from the sudden or gradual increase in terms of cooling water flow rate – variable 7 (as shown in the middle diagrams of Figure 3.5). As a result, this then affects the behaviour in the temperature condition in the tank (variable 1), where gradual reduction has been observed (as shown in the top diagrams of Figure 3.5) because the increase in cooling water flow rate. In particular, the magnitude of change for abrupt case is greater than the incipient scenario. As a result, controller 2 has responded by decreasing the signals (valve 2 – forced to open), where the valve is either fully shut (abrupt case – bottom left) or gradually closing (incipient case – bottom right) for controlling the amount of cold water fed into the heat exchanger (as shown in the bottom diagrams of Figure 3.5).

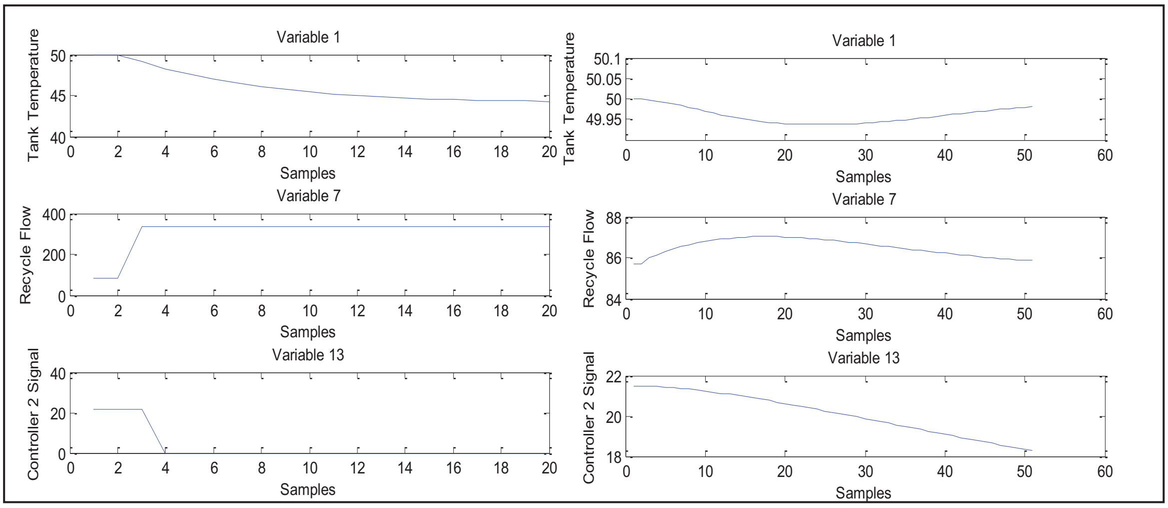

Figure 3.5: The behaviours of fault number 6 based on trends of variables 1 (top charts), 7 (middle charts) and 13 (bottom charts)

Figures 3.6 and 3.7 present the results of process monitoring with regard to F6a and F6i respectively. In analysing the overall performance, all the three PCA models developed have detected this particular fault successfully. Regarding F6a (Figure 3.6), the progressions on SPE statistics are very efficient, where the fault can be detected as fast as at 3 sampling time. The similar trend can be also observed on T^2 progressions, but the trend is not consistent, where there are few samples have returned to the normal region some time later after detection. Anyway, the overall performance has indicated abnormal signal from sample 3 until 20 as both statistics work complementary. Nonetheless, the detection time for incipient fault 6 is generally slow (Figure 3.7). In particular, the detection time for F6i based on 3, 5 and 7 PCs is 42, 36 and 23 through SPE statistics progressions. The overall trend suggests that the more PCs adopted by PCA, the detection time tends to be faster. It is also observed that the T^2 statistic has failed detecting the specified case based on the number of samples utilised. Thus, the overall detection for F6i is totally depended on the SPE outputs.

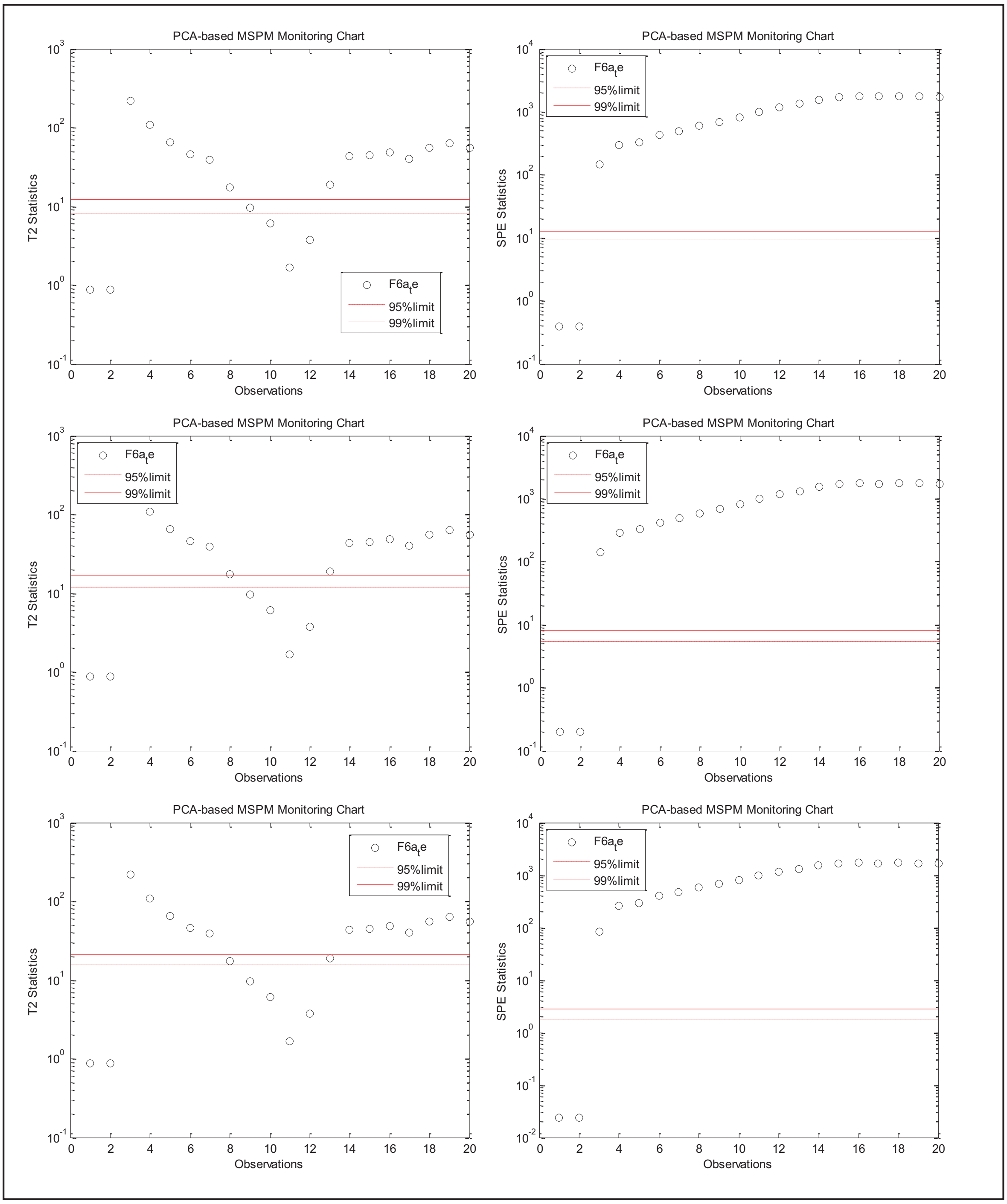

Figure 3.6: Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F6a data of CSTRwR

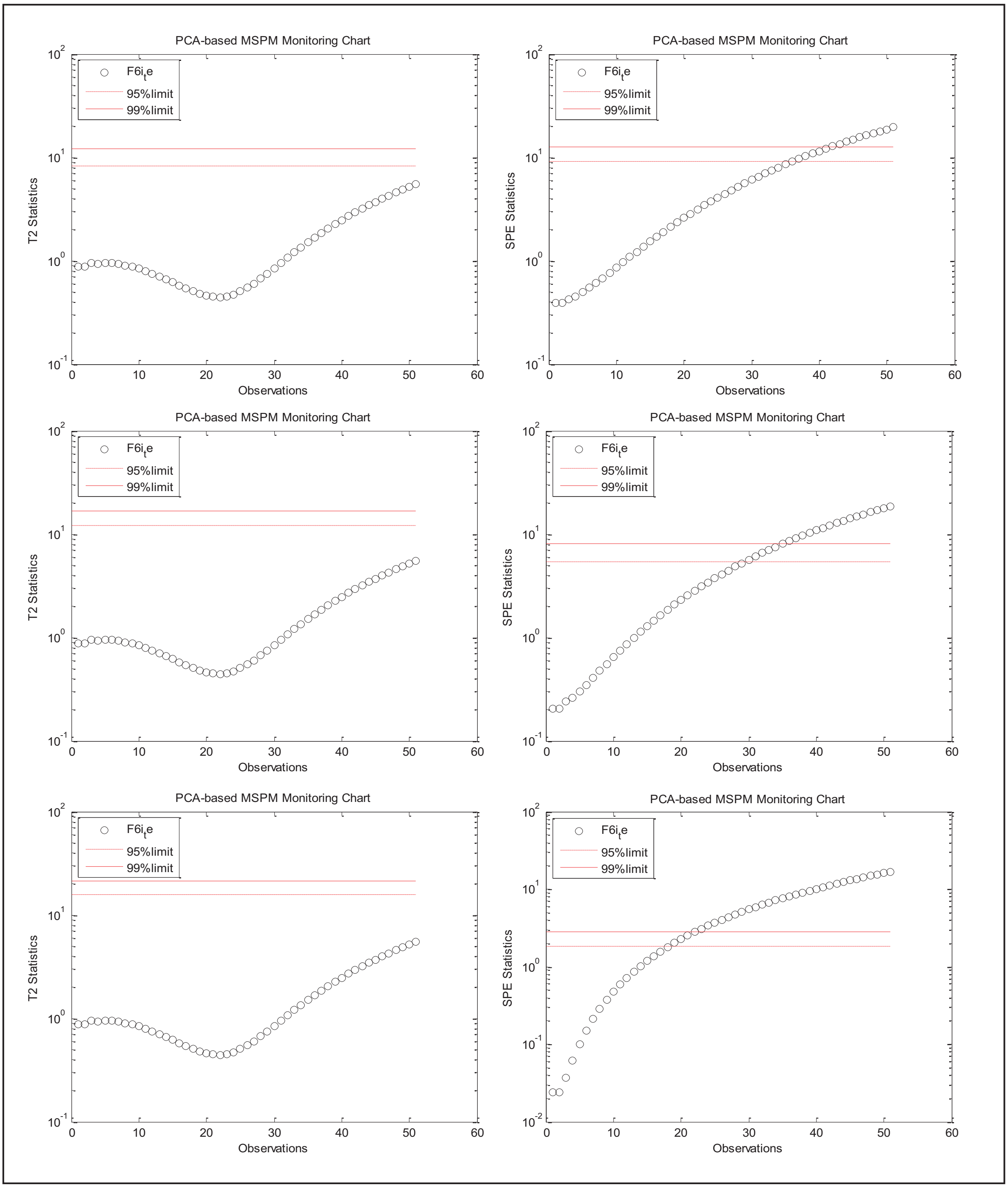

Figure 3.7: Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F6i data of CSTRwR

The second type of faults, which related to both F9a and F9i, involves blockage in the recycle streams (pipeline 4, 5 or 6 in Figure 3.1), where the impact can be clearly seen on the tank temperature condition, by which, it has indicated gradual increment in magnitude (top diagrams in Figure 3.8). This is because the amount of cold materials that recycled into the tank is decreased. In consequence of this particular changing on the tank temperature, controller 2 has responded by opening the cold water fed valve (valve 2 – forced to open) accordingly, as can be observed by the increasing signals on controller 2 (middle diagrams in Figure 3.8). At the same time, controller 3 has also acted by opening the recycle stream valve (valve 3 – forced to open) to allow more materials enter into the tank for compensating the blockage effect. This effect is very obvious on the abrupt fault case, where the controller 3 signal has increased to 100% (valve 3 is fully opened), whereas the signal is slowly increasing during controlling the incipient fault (valve 3 is opened gradually) as indicated in the bottom diagrams in Figure 3.8).

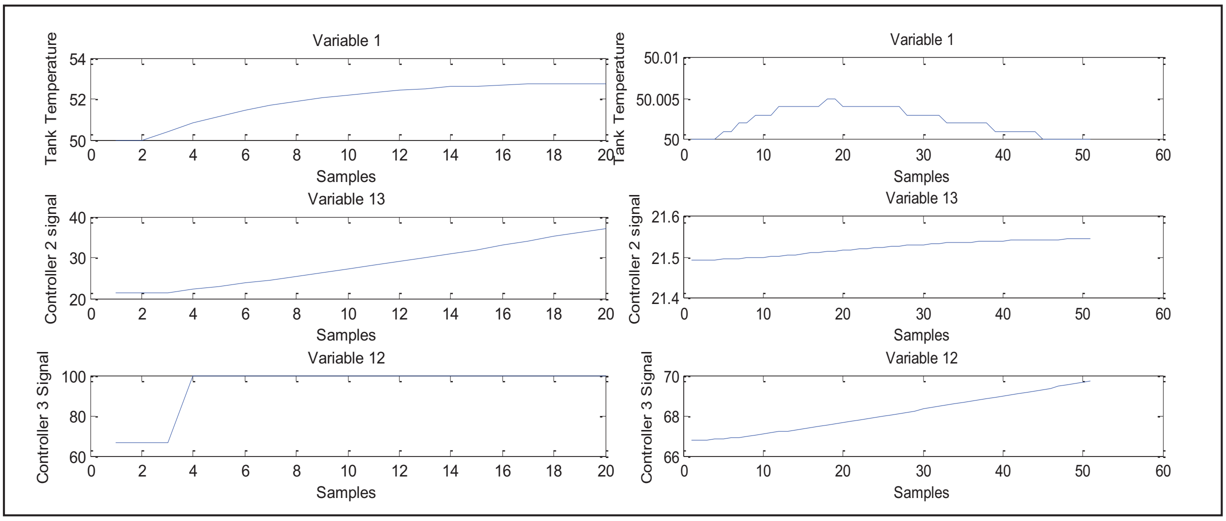

Figure 3.8: The behaviours of fault number 9 based on trends of variables 1 (top charts), 13 (middle charts) and 12 (bottom charts) of CSTRwR system for abrupt (left diagrams) and incipient (right diagrams) fault categories

Figures 3.9 and 3.10 show, respectively, the monitoring results of F9a and F9i. Figure 3.9 demonstrates that all the models have detected F9a efficiently at sampling time 3, obviously through the SPE statistic. From Figure 3.10, however, the detection time for F9i is also found generally sluggish. In particular, models with 3, 5 and 7 PCs has detected F9i at the sampling time 31, 26 and 16 respectively also solely based on SPE statistic (in this case, T^2 has once again has failed in detection as similar to the F6i case previously).

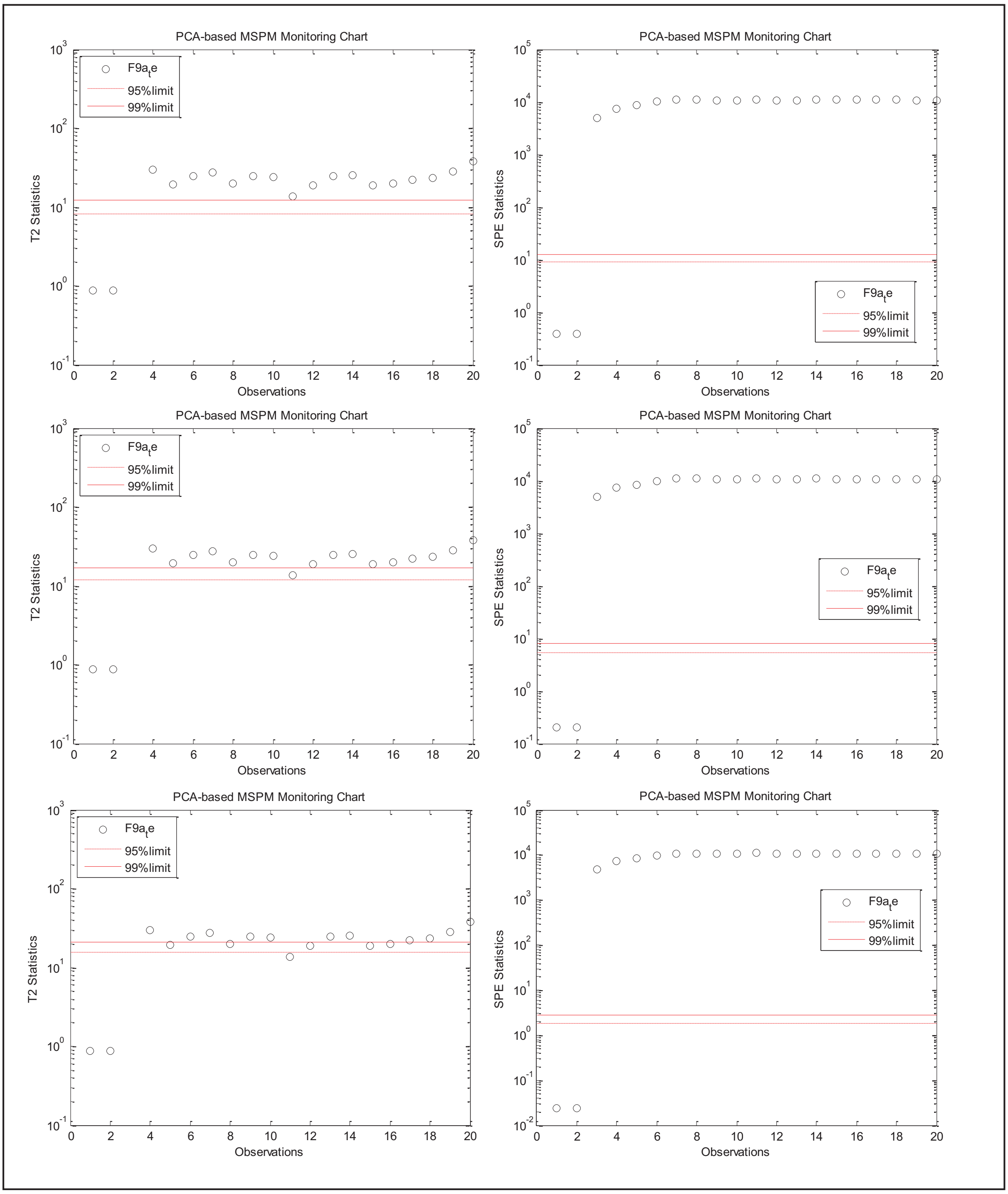

Figure 3.9: Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F9a data of CSTRwR

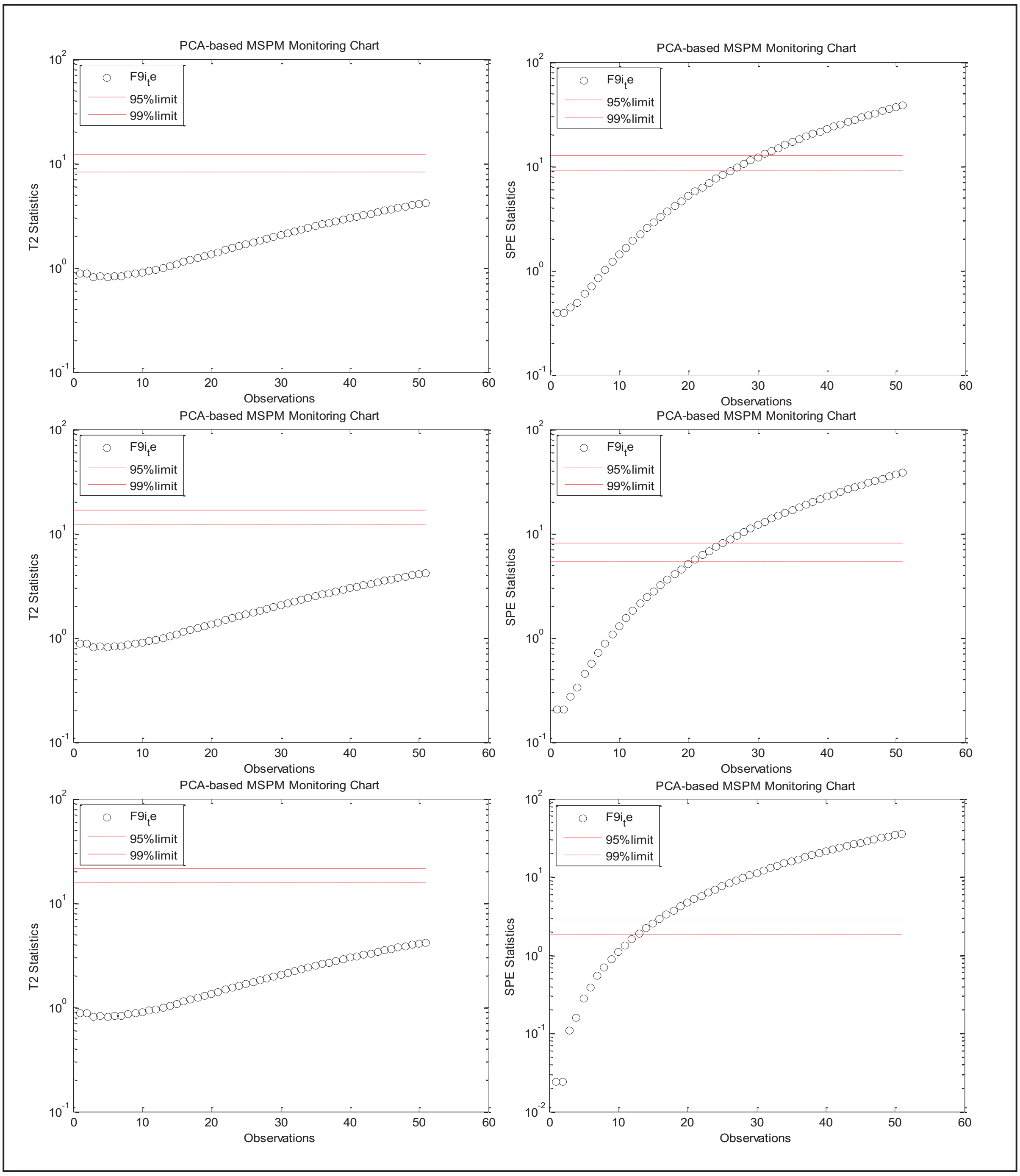

Figure 3.10: Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F9i data of CSTRwR

Lastly, Figure 3.11 presents the nature of the third fault behaviour that associated to F11a and F11i. The graph shows that there is a sudden or gradual decrease in magnitude in the concentration of input stream (variable 9) as indicated in the top diagrams in Figure 3.11. The change is then affecting on the product concentration (variable 8) by denoting decreasing trending as shown in the bottom diagrams in Figure 3.11.

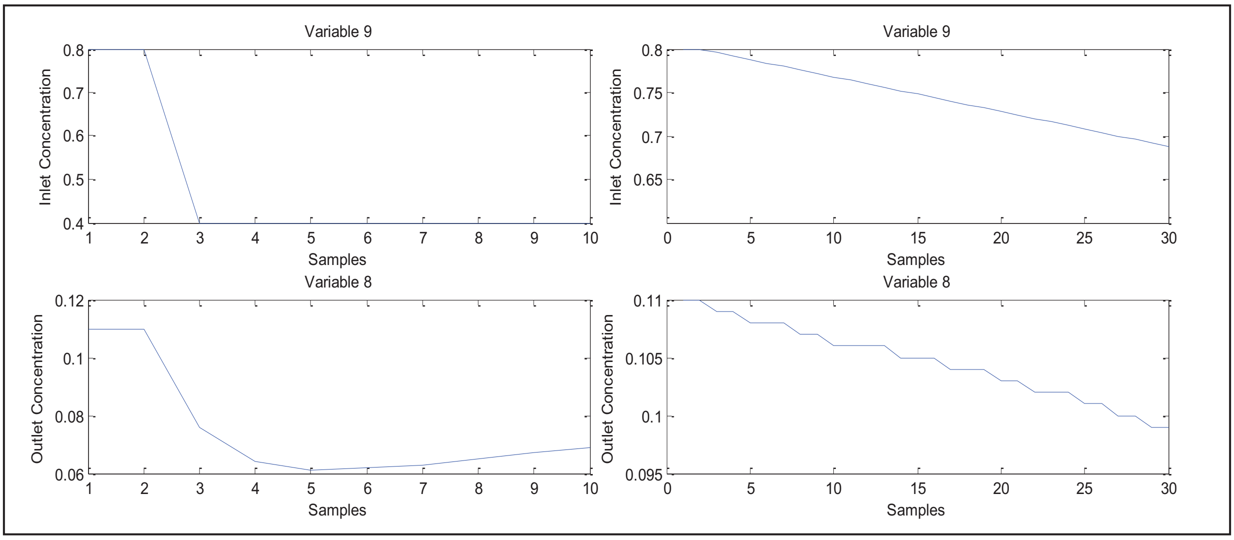

Figure 3.11: The behaviours of fault number 11 based on trends of variables 9 (top charts), and 8 (bottom charts) of CSTRwR system for abrupt (left diagrams) and incipient (right diagrams) fault categories

The monitoring performances with respect to F11a and F11i are shown in Figure 3.12 and 3.13 respectively. From Figure 3.12, both of the statistics have managed detected F11a efficiently at sampling time 3. Figure 3.13, on the other hand, has depicted that F11i can be detected much faster compared to the previous cases. In particular, the detection time is found to be at sampling time 5 for PCA models 3 and 5, whereas PCA model 7 has detected slightly later, which is at sampling time 11 through SPE statistic. The T^2 progressions also have shown detections but in much slower period, regardless of the models that applied.

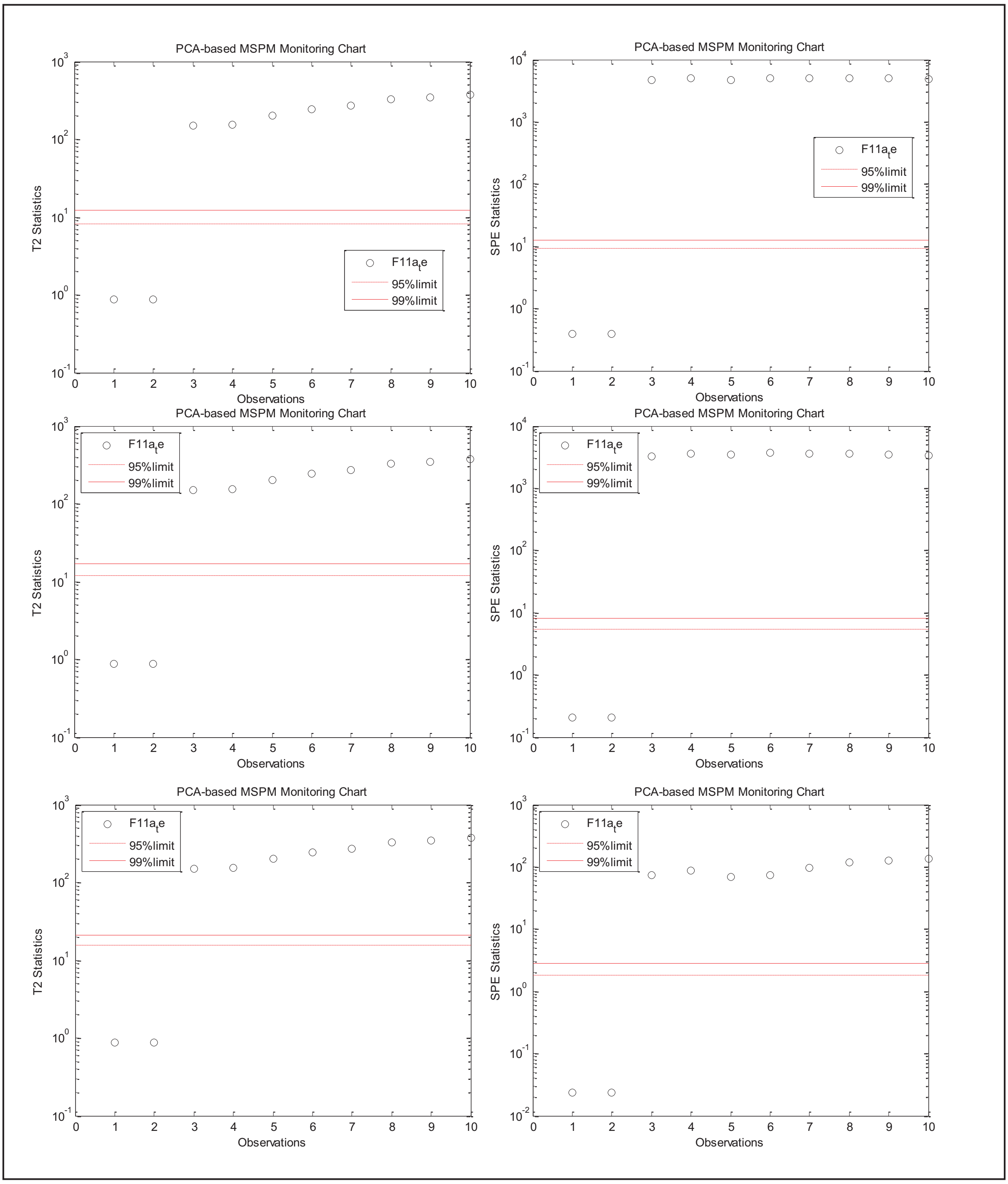

Figure 3.12: Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F11a data of CSTRwR

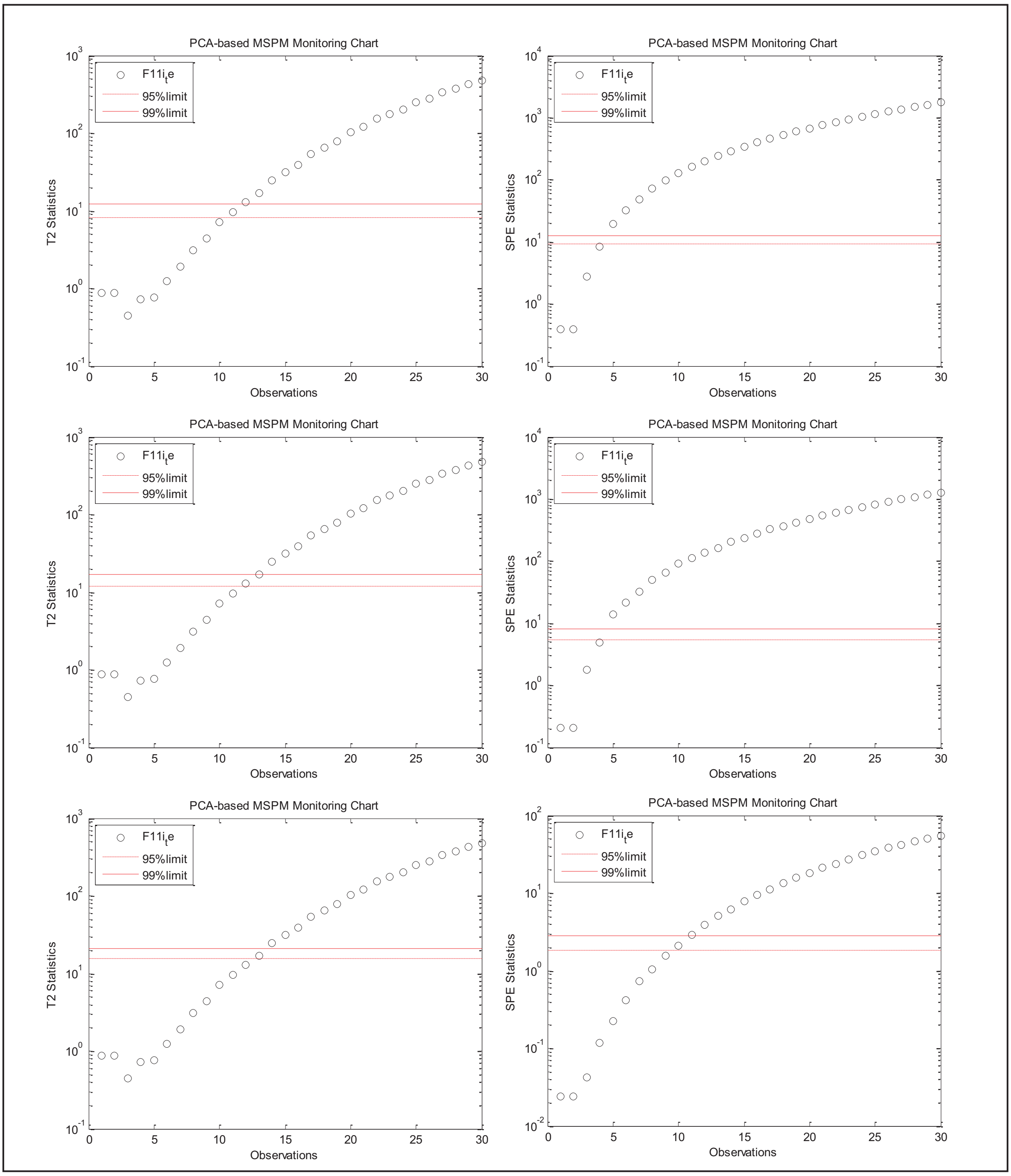

Figure 3.13: Progressions of T^2 (left) and SPE (right) of PCA models with 3 PCs (top), 5 PCs (middle), and 7 PCs (bottom) for F11i data of CSTRwR

3.3 Summary

A CSTRwR system was utilised as the case study for monitoring evaluation. In general, this particular unit operation represents a system where a set of variables are monitored and the impact of the disturbances on the process is quite observable as well as the trend persists over the time of operation. Besides, this chapter has also discussed the PCA models which are developed according to the original NOC data obtained from simulation. In addition, the monitoring limits are also verified to be reliable for monitoring (tested through the second set of NOC data). This chapter also demonstrates various fault detection performances using the PCA technique. The generic trend shows that PCA is very efficient in detecting any of abrupt fault cases, whereas the same models may have some difficulties (slower detection) in dealing with the incipient faults. All these results will be used against the outcomes of the proposed CMDS systems for evaluation. The following subsequent chapters focus on the comparative analysis of monitoring performances between the proposed methods and the linear PCA based monitoring systems developed in this chapter. As the proposed frameworks are divided into three main bases, therefore, each chapter will be dedicated specifically to one of these generic procedures respectively.

CHAPTER 4

FRAMEWORK I: MDS-BASED MSPM SYSTEM USING MOVING WINDOW CMDS PROJECTION

4.1 Introduction

This chapter presents the methodology and results of the first framework, which is based on the standard CMDS procedures (sCMDS-MSPM). This study assumes that, whenever a fault situation occurred, those faulty variables have the tendency of freeing themselves from the normal cluster. Therefore, this method is mainly trying to identify those behaviors by the use of two proposed monitoring statistics. In particular, the execution can be performed by monitoring the movement of individual variables from the global origin of the scores as well as measuring the changes in terms of inter-distances magnitude among of the variables. The true advantage of CMDS as opposed to PCA in this application actually lies on its ability to preserve the association among the objects (or 'variables' in this context) within the new reduced dimensional space, where it is strictly configured and mapped based on the original dissimilarity scales. Thus, analysis on the scores from the other non-selected dimensions can be avoided (as normally performed through SPE). Besides, this unique feature will in turn, perhaps support the system to be effectively implemented in less dimensionality against the linear PCA for monitoring.

The remainder of this chapter is divided into three sections. Section 4.2 describes the proposed CMDS based monitoring procedures. The results of the proposed approach on the case studies, as well as its critical reflection, are presented in Section 4.3. Lastly, a summary is given in Section 4.4.

4.2 Methodology

The sCMDS-MSPM procedures are generally divided into two main phases – phases I and II as depicted in Figure 4.1. From Figure 4.1, the first phase relates to the model development using normal operating condition (NOC) data (off-line modelling) whereas the second phase facilitates for monitoring of the new process data (on-line monitoring).

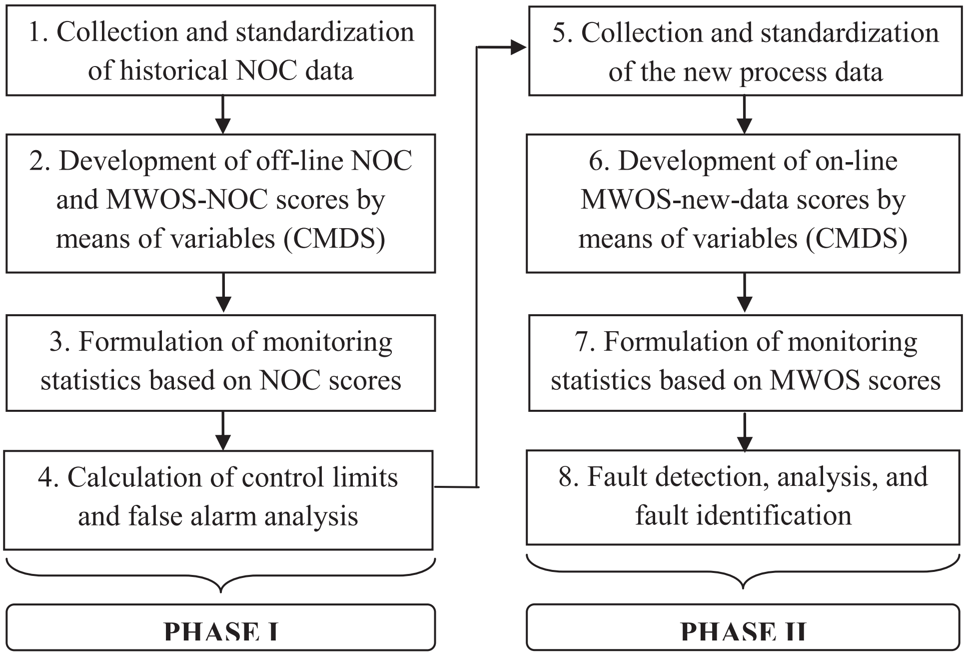

Figure 4.1. CMDS-based MSPM framework

4.2.1 Phase I Procedures

The following discussions are related to steps 1 to 4 in Figure 4.1. Firstly (**1st step**), a set of NOC data, $\mathbf{X}_{m \times n}$ (*m*: variables, *n*: samples), are identified off-line based on the historical process data archive. The data are then standardized to zero mean and unit variance (because the data involves variables with various units). In contrast to the first step of Figure 2.1, the NOC data is then divided into two sets (after standardization), where the first ($\mathbf{X}_{\mathbf{NOC1}}$) will be used in developing the optimized NOC scores and the second ($\mathbf{X}_{\mathbf{NOC2}}$) will be integrated with $\mathbf{X}_{\mathbf{NOC1}}$ for identifying the monitoring limits. This is important as the variables' scores structure cannot be directly used in providing the distribution model information.

In the **2nd step**, the first set of the NOC data is then compressed and converted into a set of variable score configuration in the reduced dimensional domain by using the standard CMDS procedures as discussed in Chapter 2. The developed CMDS-NOC scores, $\mathbf{Y}_{\mathbf{NOC}}$, will be in the form of *m* by *p* (*m*: number of variables; *p*: number of compressed dimensions) and assumed (with certain degree of proximity – equation 2.27) as representing the original correlations among of the monitored variables. When integrating to $\mathbf{X}_{\mathbf{NOC2}}$ (for the projection of the other NOC scores), a moving-window mechanism is applied and known as moving-window-observation sample of NOC (**X**$\mathbf{X}_{\mathbf{MWOS\text{-}NOC}}$). In particular, this mechanism is operated such that the newly measured sample is added to the data frame by taking the oldest sample out from the data window. In this way, the size of the $\mathbf{X}_{\mathbf{MWOS\text{-}NOC}}$ matrix will be maintained at *m* by *n* over the time, especially when a new sample becomes available. Later, the standard CMDS procedures are applied to develop the scores for $\mathbf{X}_{\mathbf{MWOS\text{-}NOC}}$, which is $\mathbf{Y}_{\mathbf{MWOS\text{-}NOC}}$ (in the later discussions the term $\mathbf{Y}_{\mathbf{sCMDS}}$ will be used instead of $\mathbf{Y}_{\mathbf{MWOS\text{-}NOC}}$ as to generalize the equation application especially when considering all the other faulty operation samples).

The 3rd **step** basically involves computing the monitoring statistics. The first parameter (S_{m1}) is shown mathematically in Equation 4.1 as well as illustrated in Figure 4.2.

Statistic 1 for magnitude of deviation (S_{m1}): Changes in Euclidean distance from the global origin.

$$S_{m1} = \sum_{i=1}^{m} \left(\left\{ \sum_{j=1}^{p} \left[y_{sCMDS}(i,j) \right]^2 \right\}^{0.5} - \left\{ \sum_{j=1}^{p} \left[y_{NOC1}(i,j) \right]^2 \right\}^{0.5} \right)^2 \tag{4.1}$$

where i is the variable index, j is the dimension index, p is the number of PCs, y_{sCMDS} represents MWOS scores coordinates , and y_{NOC1} represents the original NOC scores coordinates.

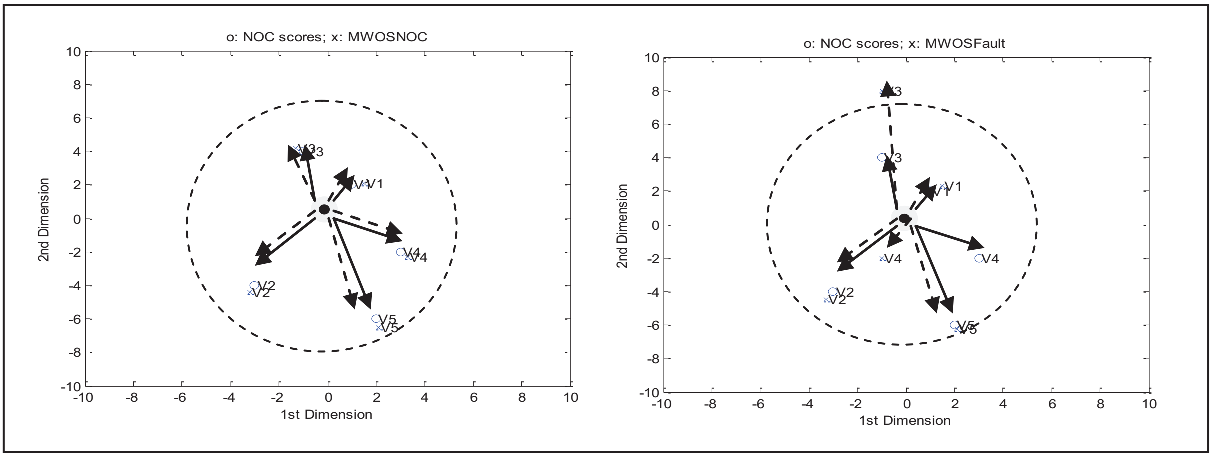

Figure 4.2: Illustration of S_{m1} based on the plots of NOC scores vs MWOS-NOC scores (left diagram); NOC scores vs MWOS-fault scores (right diagram)

It generally relates to the sum of squared errors in terms of Euclidean distance of each score from the global origin between the $\mathbf{Y}_{\mathbf{sCMDS}}$ and $\mathbf{Y}_{\mathbf{NOC1}}$ configurations. Yunus and Zhang (2010a) mentioned that this measure for symbolizing the SPE instead of T^2 because it has similarity with SPE in terms of mathematical formulation structure, nonetheless, it was later realized that the parameter is strongly connected to the measure of distance from a single reference point rather than on variable association information.

In each of the plots in Figure 4.2, an imaginary boundary (dashed-circle) can be drawn as representing the boundary of the normal score cluster. The left diagram shows that both score configurations (original NOC and MWOS-NOC) are very close together, thus, depicting the normal operation behaviour. Nonetheless, whenever a fault occurred (right diagram), those faulty variables will either move drastically or gradually away from the normal cluster, as shown by 'V3', and this trend can be easily picked up by S_{m1}.

The objective of the second statistic (S_r), however, is to measure the consistency of the current variables' association according to the specified original NOC model. This can be easily executed by way of measuring the sum of squared errors in terms of dissimilarity measures between the $\mathbf{Y}_{\mathbf{sCMDS}}$ and $\mathbf{Y}_{\mathbf{NOC1}}$ scores coordinates as defined in Equation 4.2.

Statistic for relationship (S_r): Sum of squared errors in terms of dissimilarity measures.

$$S_r = \sum\left(\sum_{i=1}^{m}\sum_{j=1}^{m}\left[d_{sCMDS(ij)} - \delta_{NOC1(ij)}\right]^2\right) \; ; i \neq j \tag{4.2}$$

where i and j are respectively the index for rows and columns of the dissimilarity matrix, m is the number of variables, d_{sCMDS} is the dissimilarity matrix of the MWOS scores, and δ_{NOC1} is the dissimilarity matrix of the first NOC scores.

The same configuration of Figure 4.2 is used again in illustrating the idea of S_r as shown in Figure 4.3, where distances from variable 'V2' to other variables are shown. From the right diagram of Figure 4.3, when 'V3' is deviated from the normal cluster, the impact on S_r can be obviously seen by indicating a relatively large distance from 'V2' compared to the same original measure as indicated in the left diagram of Figure 4.3. In another instance, the abnormal condition shown by 'V4' in the right diagram of Figure 4.3, can also be detected by S_r, where the distance between 'V4' and 'V2' is comparatively smaller than the same original measure depicted in the left diagram in Figure 4.3. These trends show that either 'V3' or 'V4' is contributing or being affected by the fault.

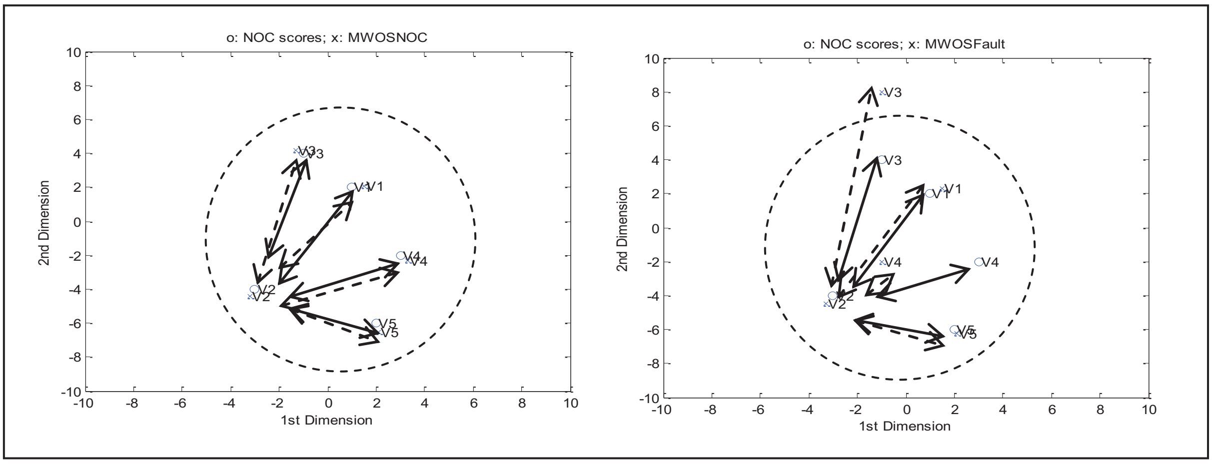

Figure 4.3: Illustration of S_r based on the plots of NOC scores vs MWOS-NOC scores (left diagram); NOC scores vs MWOS-fault scores (right diagram)

All of these monitoring statistics can be assumed following the chi-squared distribution as discussed in Nomikos and MacGregor, (1995) because they are all using the similar equation structure which is based on 'sum of squared errors' and the "error" terms can be reasonably assumed to follow normal distribution. The monitoring limits for each of the statistics are calculated using Equation 4.3 (**4th step**).

$$\lim_{\alpha} = (\nu / 2\overline{m})\chi^2_{2m^2/\nu,\alpha} \tag{4.3}$$

where α equals to 0.05 and 0.01 for warning (95%) and control (99%) limits respectively, $\overline{m}$ and v are respectively representing the means and variances for each of the statistics.

Once all of these procedures are carried out, false alarm rate (FAR) analysis is conducted to evaluate the robustness of the monitoring statistics (Chiang et al., 2001). Equation 4.4 is used to obtain the FAR results where the aim is to get a very low FAR. High FAR indicates that the developed limit settings are unsuitable to be used for monitoring, and therefore, another approach should be initiated to optimize the situation. Hence, the window size settings are utilized for the optimization purpose. In other words, different window sizes will be implemented for each of the algorithms to vary the monitoring outputs.

$$FAR = \frac{\text{Total amount of statistics in the NOC set located beyond the 99\% limit}}{\text{Total amount of NOC samples}} \quad (4.4)$$

4.2.2 Phase II Procedures

In the procedures of on-line monitoring, which are pertaining to steps 5 to 8 in Figure 4.1, on-line measured data are collected and standardised using the means and standard deviations for the NOC data and subsequently transformed into the MWOS structures as described in step 2 (**5th step**). The multivariate scores (**6th step**) as well as monitoring statistics (**7th step**) are computed following the same procedures described in the previous section. Finally, real time monitoring is performed by observing the progression of all monitoring statistics on the Shewhart-type multivariate monitoring chart (**8th step**). A fault is directly detected whenever one of the monitoring statistics exceeds the specified 99% monitoring limit based on a consecutive number of samples.

A systematic as well as comprehensive analysis is conducted to compare the monitoring results of sCMDS and linear PCA based monitoring systems. The following performance indicators are compared, the total number of detected cases and total number of fastest detection cases, which are both depended on the fault detection time (FDT) outputs. Both of these performance factors are important as to identify the credibility of the individual system in tackling various abnormal cases. In short, if the proposed system can detect equivalently or closely by way of the number of cases relative to the conventional method, it basically means that the proposed framework is practically working.

On the other hand, FDT can be regarded as the number of samples between the time a fault was introduced in the process and a monitoring index (either S_{m1} or S_r) exceeded the 99% control limit for the first time (subject to the pre-specified condition, in terms of consecutive number of samples, that defining the fault detection execution). Thus, the term is actually referring to a measure of rapidity of the system in detecting the fault or promptness according to Chiang et al., (2001). There is also another indicator that can be utilised for analyzing the monitoring performance, which is known as missed detection rate (Chiang, et al., 2001). This measure is defined as the ratio of faulty samples being detected and the total number of faulty samples. Nonetheless, this study believes that this particular parameter shares the similar conceptual objective as with the FDT, which is to identify the sensitivity of the monitoring system in detecting the faults, but both are different in terms of mechanisms. This measure is importantly required however, especially when considering the dynamic trends of the monitoring progressions. Thus, if the proposed methods can achieve quicker detection, it basically means that the assumed framework has the potential to be subjectively efficient compared to the traditional system.

There are also other performance factors that can be utilized for analysis, particularly in comparing between the monitoring system performances. As an example, Venkatasubramaniam et al., (2003) have proposed ten characteristics that are desirable for any process monitoring which include quantitative and qualitative measures. Among those criteria, quick detection (short FDT) and robustness (less FAR) have been particularly addressed in this study, and both are quantifiable. While others seemed to be either more on diagnosis oriented such as isolable capability, classification error estimate and explanation facility or simply very subjective in definition such as those concerning modelling, storage and computational requirements. There are also other factors which are relevant for consideration such as novelty identification, adaptability in dealing various operation modes and also multiple fault diagnostic capability. However, these three factor are more suitably to be considered in the next phase of advanced monitoring, due to the nature of complexity as well as constraints of the environment that those factors are emphasizing and struggling to solve (which are contextually different from the issues focused in this work). Therefore, the two monitoring performance factors together with FAR are the main outcomes that sufficient for evaluation corresponding to the specified objectives which have been set in this study.

The contribution plot technique is proposed to identify the potential variables that possibly connected to the detected problem as given in Equation (4.5).

$$\left(CMDS\ Statistics\right)_j = \sum_{i=1}^{m}\left(X_i\right)_j \tag{4.5}$$

Where, $(CMDS\ Statistics)_j$ = CMDS statistics (S_{m1} or S_r) at a particular sampling time '*j*'.

$(X_i)_j$ = contribution of the *i*th variable to CMDS statistics at a particular sampling time '*j*'.

Equation (4.5) is applicable to both statistics proposed. Thus, once the system has detected a fault, fault identification will be conducted by using contribution plot approach on the corresponding statistic or statistics that detected the fault. This is performed iteratively from the instance of first detection until a definite period of time, whereby the trend of deviations persists permanently.

4.3 Results and Analysis

This section presents the performances of CMDS based monitoring system on various faults and compares with those from a a linear PCA based monitoring system. In particular, FAR, the total number of cases detected, and the total number of fastest detection cases of both monitoring systems are compared.

The same sets of NOC data which have been used in developing the PCA models, was used in constructing the sCMDS NOC model. As described in the previous section, the NOC data were divided into two groups, $\mathbf{X_{NOC1}}$ and $\mathbf{X_{NOC2}}$ (each contains 50 samples). The first ($\mathbf{X_{NOC1}}$) was used in developing the NOC scores and the second ($\mathbf{X_{NOC2}}$) were integrated with $\mathbf{X_{NOC1}}$ for identifying the monitoring limits. The result on the accumulated portion of data variance explained by the dimensions (pertaining to $\mathbf{X_{NOC1}}$) is given in Figure 4.4.

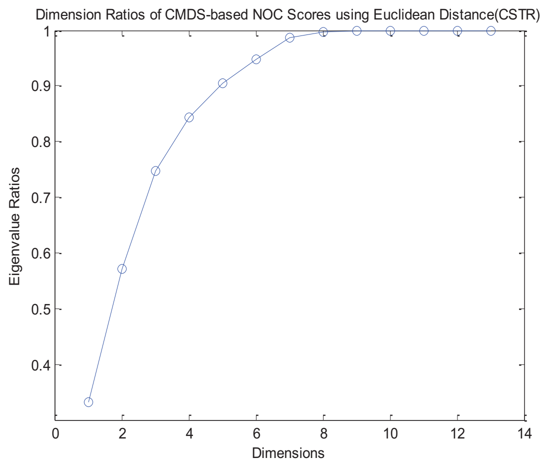

Figure 4.4. Accumulated portion of data variation explained by the dimensions of the CMDS model of $\mathbf{X_{NOC1}}$

Figure 4.4 indicates that the MDS model with 3, 5 and 7 dimensions represent around 75%, 90% 99% of the data variation respectively, which are similar to the PCA models.

Nonetheless, the interpretation on these ratios is different. From the CMDS point of view, the ratios provide the sense of proximity of the reconstructed scores with respect to the measures of dissimilarity scales (in terms of variable inter-distance measures) instead of variances. It basically means that, the model representing 90% of data variation is perceived to be more accurate compared to the one representing 75% of data variation as far as the original dissimilarity scales are concerned. Therefore, this condition cannot be used to justify concretely that the higher percentage model should perform better in detecting faults in relative to the smaller dimensional model. However, those three dimension settings were used in this study to analyze and compare their performance.

The overall results of FAR analysis on all the cCMDS methods have shown zeros rates. This indicates that the established monitoring limits should be robust in the monitoring operation. The results are also comparable to those of linear PCA based monitoring systems presented in Chapter 3.

Tables 4.1 and 4.2, on the other hand, present the overall monitoring performances in terms of fault detection time (FDT) of both sCMDS and PCA systems based on 3, 5 and 7 dimensions/PCs settings corresponding to abrupt and incipient fault cases respectively. In all the cases, the faults were detected such that at least 3 successive samples of either monitoring statistics were located above their 99% control limits. In both tables, the fastest detection time is indicated in bold.

From Table 4.1, the overall performances have shown that all the PCA models have managed detecting all of the abrupt fault cases. This similar performance is also shared by the sCMDS method which particularly applies the Euclidean distance regardless of the dimensions and window sizes used. The results on using City-block distance is also productive (100% detection based on dimension settings of 3 and 7 and adopting window size 15), even though the other application specifications have demonstrated slightly lower number of detections. In addition, all of the PCA models have denoted that they are both effective as well as efficient, where all the abrupt fault cases can be detected as fast as at 3 samples (except case 7 based on dimensions 3 and 5). Regarding sCMDS, the best performance is produced by methods that implementing Euclidean distance and window size 15 (altogether 8 cases with fast detection) for all dimensions that applied. In analysing the results based on those 3 cases where sCMDS based systems did not produce fastest detections, the FDT based on sCMDS is not much longer compared to PCA. The generic City-block results, on the other hand, show low number of fastest detection cases, nonetheless, the delayed FDT of City-block in relative to the PCA and Euclidean outcomes are minor. Thus, the overall performance of sCMDS can be perceived as comparable to linear PCA.

Table 4.2 shows that the monitoring systems based on PCA and sCMDS with Euclidean distance have successfully detected all the incipient fault cases that investigated (100 % detection) consistently regardless of the application settings. As similar to the previous discussion on abrupt faults, the performance of sCMDS based monitoring system with City-block distance is conditional, where the models with 5 and 7 dimensions have shown 100% detection with window size 15, whereas others have produced 10 cases detected (which is only one case undetected). However, the significant impact of using sCMDS can be observed particularly based on the total number of fastest detection case, where the best performance is all dominated by sCMDS with Euclidean distance and window size 5 for every dimension cosidered. In particular, 8 out of 11 cases that are detected are found to be the fastest detection, which is hugely different from those cases of PCA as well as City-block. In analysing the results in detail, majority of the cases have shown that the difference in FDT between the monitoring systems based on sCMDS with Euclidean distance and PCA are found very large. This observation suggests that sCMDS has an important advantage over linear PCA particularly in dealing with slow or incipient faults.

Table 4.1: Fault detection times of monitoring systems based on sCMDS and PCA for abrupt fault cases

FAULT DETECTION TIME (samples)																						
Fault Cases	Dimensions	3							5							7						
	Methods	sCMDS						PCA	sCMDS						PCA	sCMDS						PCA
	Window sizes	5		10		15			5		10		15			5		10		15		
	Scales	E	C	E	C	E	C		E	C	E	C	E	C		E	C	E	C	E	C	
1a	Fault Detection Time	2	ND	2	ND	**1**	9	**1**	2	ND	2	ND	**1**	11	**1**	3	ND	2	ND	**1**	9	**1**
2a		3	ND	3	ND	2	14	**1**	5	ND	4	ND	2	ND	**1**	5	ND	4	ND	2	14	**1**
3a		**1**	2	**1**	**1**	**1**	**1**	**1**	**1**	2	**1**	2	**1**	**1**	**1**	**1**	2	**1**	2	**1**	**1**	**1**
4a		**1**	7	**1**	6	**1**	5	**1**	**1**	7	**1**	6	**1**	6	**1**	**1**	7	**1**	6	**1**	5	**1**
5a		**1**	5	**1**	4	**1**	4	**1**	**1**	5	**1**	5	**1**	4	**1**	**1**	5	**1**	5	**1**	4	**1**
6a		**1**	4	**1**	4	**1**	3	**1**	**1**	4	**1**	4	**1**	3	**1**	**1**	5	**1**	4	**1**	3	**1**
7a		8	ND	8	15	8	13	**5**	8	ND	8	15	7	14	**2**	11	ND	8	ND	8	13	**1**
8a		3	11	3	9	2	8	**1**	4	11	3	10	2	9	**1**	4	11	3	10	2	8	**1**
9a		**1**	2	**1**	**1**	**1**	**1**	**1**	**1**	2	**1**	2	**1**	**1**	**1**	**1**	2	**1**	2	**1**	**1**	**1**
10a		**1**	3	**1**	3	**1**	2	**1**	**1**	3	**1**	3	**1**	2	**1**	**1**	3	**1**	3	**1**	2	**1**
11a		**1**	2	**1**	2	**1**	**1**	**1**	**1**	2	**1**	2	**1**	2	**1**	**1**	2	**1**	2	**1**	1	**1**
Total number of detected cases		**11**	8	**11**	9	**11**	**11**	**11**	**11**	8	**11**	9	**11**	10	**11**	**11**	8	**11**	8	**11**	**11**	**11**
Total number of fastest detected cases		7	0	7	2	8	3	**11**	7	0	7	0	8	2	**11**	7	0	7	0	8	2	**11**
Legends : E=Euclidean C=City-block ND=no detection																						

Table 4.2: Fault detection times of monitoring systems based on sCMDS and PCA for incipient fault cases

FAULT DETECTION TIME (samples)																						
Fault Cases	Dimensions	3							5							7						
	Methods	sCMDS						PCA	sCMDS						PCA	sCMDS						PCA
	Window sizes	5		10		15			5		10		15			5		10		15		
	Scales	E	C	E	C	E	C		E	C	E	C	E	C		E	C	E	C	E	C	
1i	Fault Detection Time	3	9	4	8	3	8	**2**	**4**	11	**4**	9	**4**	8	**4**	**4**	11	**4**	9	**4**	8	5
2i		7	ND	7	13	7	12	**3**	8	ND	**7**	14	**7**	13	**7**	8	ND	**7**	15	**7**	12	8
3i		4	14	4	10	4	9	**2**	3	15	4	10	3	10	**2**	3	15	4	11	3	10	**1**
4i		**3**	**3**	8	8	24	14	29	**3**	**3**	6	7	11	11	25	3	**2**	6	6	9	14	15
5i		**3**	**3**	16	ND	16	ND	10	**3**	**3**	8	8	12	ND	9	3	**2**	7	8	12	ND	22
6i		**3**	**3**	8	8	13	12	40	**3**	**3**	6	7	9	10	34	**3**	**3**	5	6	9	12	21
7i		**3**	**3**	8	8	13	12	39	**3**	**3**	6	6	9	10	32	3	**2**	5	6	9	12	21
8i		**3**	**3**	8	8	13	12	27	**3**	**3**	6	6	9	10	24	3	**2**	5	6	9	12	15
9i		**3**	**3**	8	8	13	12	29	**3**	**3**	6	7	9	11	24	3	**2**	5	7	9	12	14
10i		**3**	**3**	8	8	13	12	30	**3**	**3**	6	7	9	10	25	3	**2**	5	6	9	12	16
11i		**3**	20	5	11	5	10	3	**3**	23	6	13	6	11	3	**3**	**3**	6	13	6	10	9
Total number of detected cases		**11**	10	**11**	10	**11**	10	**11**	**11**	10	**11**	**11**	**11**	10	**11**	**11**	10	**11**	**11**	**11**	10	**11**
Total number of fastest detected cases		**8**	**8**	0	0	0	0	3	**8**	**8**	2	0	2	0	3	2	**8**	2	0	2	0	1
Legends : E=Euclidean C=City-block ND=no detection																						

Those cases that discussed individually in Chapter 3 are utilised again for illustration in this chapter. Figures 4.5 and 4.6 present the monitoring statistic progressions for F6a and F6i respectively. From Figure 4.5, it can be seen that the overall results on F6a demonstrated that both statistics are effective as well as efficient, where this particular fault can be detected as early as at sampling time 3. In compared to PCA, both methods are found to have equal performance based on this particular case. With regard to F6i, however, Figure 4.6 has depicted that both statistics are significantly productive, where this particular fault can be efficiently detected at sampling time 5. This is significant improvement compared to PCA, where the FDTs are found to be generally more than 20 sampling time as shown in Chapter 3.

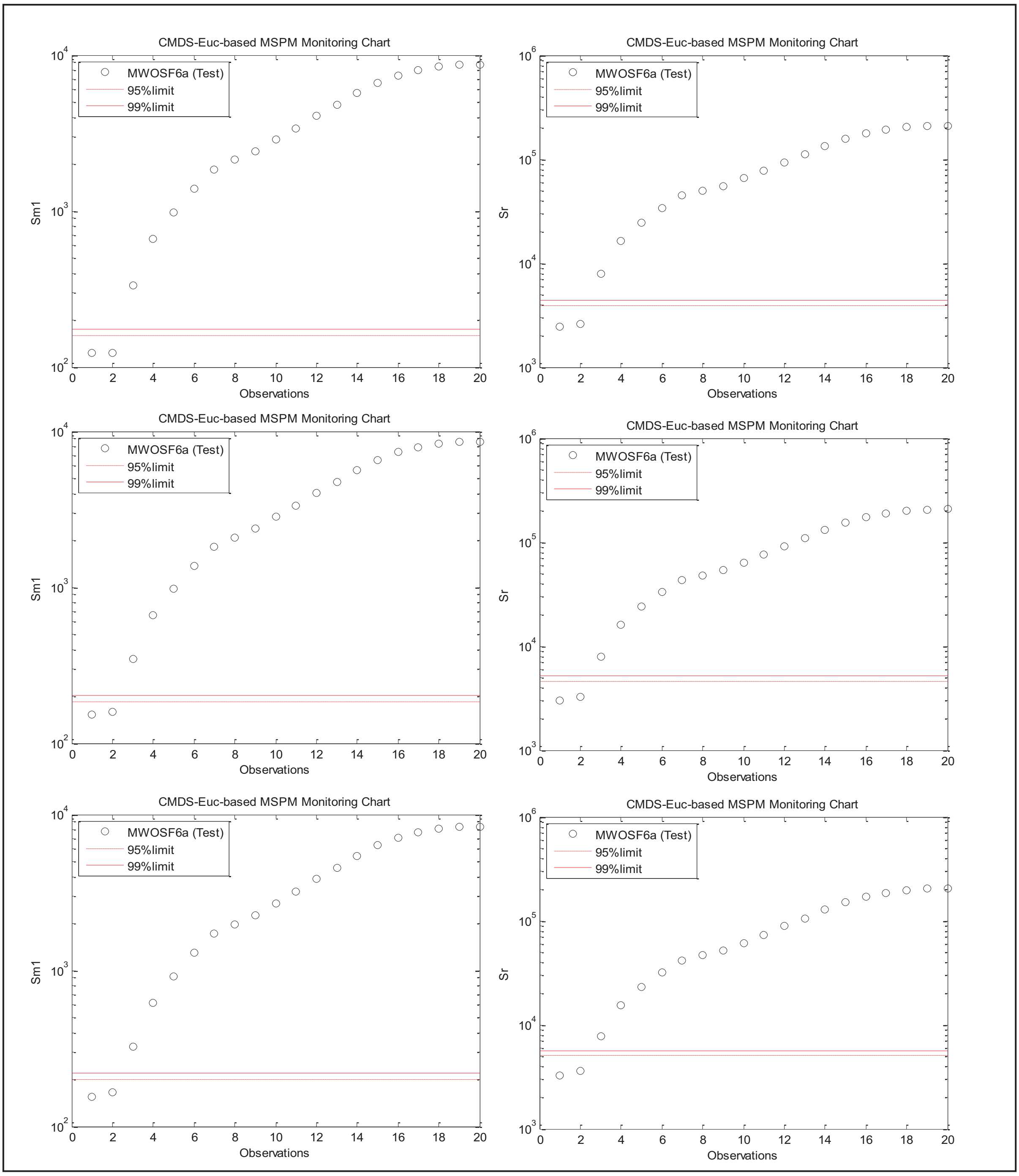

Figure 4.5: Monitoring progression of S_{m1}(left) and S_r (right) on F6a based on sCMDS models using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

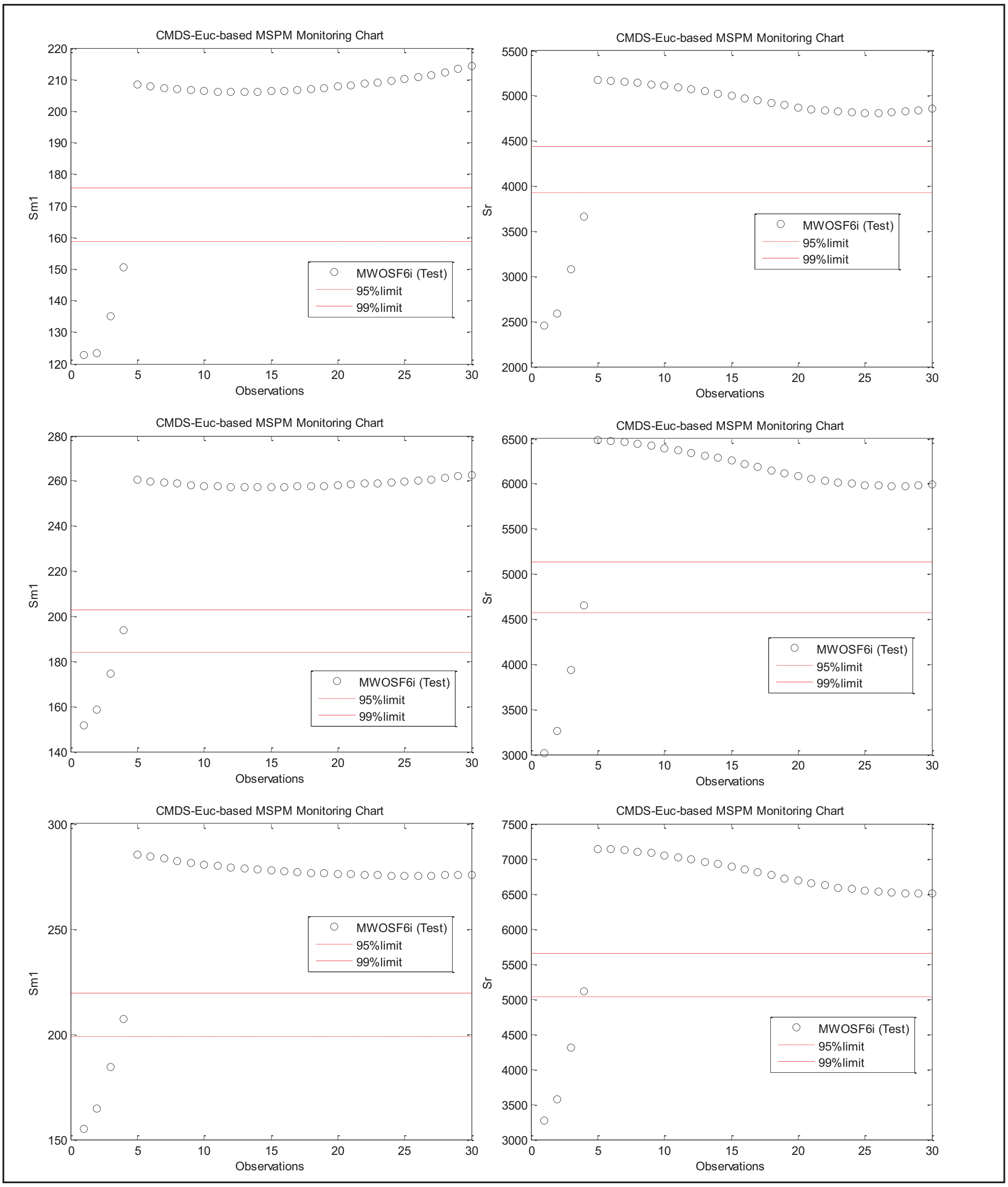

Figure 4.6: Monitoring progression of S_{m1}(left) and S_r (right) on F6i based on sCMDS models using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

Fault identification is carried out through contribution analysis. Figures 4.7 and 4.8 shows the corresponding contribution plots of sCMDS based monitoring system with Euclidean distance and window size 5 for F6a and F6i respectively. In particular, all contribution plots in Figure 4.7 indicate that variables 7 and 13 have the most significant contributions to the monitoring statistics under F6a at sampling time 10. Prior to this, variable 7 has been seen as the main sole contributor to the fault particularly from sampling time 3 until 9, whereas both variables 7 and 13 have been sustained to be strongly connected to the fault from sampling time 10. This fault leads to an increase in the cooling water flow rate (variable 7) which results in a reduction in reactor temperature. The reactor temperature control system then attempts to reduce the cooling water flow rate in order to maintain the reactor temperature by reducing controller 2 output (variable 13).

Meanwhile, the contribution plots in Figure 4.8 have indicated variable 13, as the only main contributor to F6i. This can be observed clearly through S_{m1} (top diagrams in Figure 4.8) at sampling time 20 and this is consistent until to the last sample (the results based on earlier sampling time have shown no clear trending). Unfortunately, the progression of S_r cannot identify the specific trend clearly based on the similar period of time that taken by S_{m1}. It is also interesting to observe that variable 7 did not importantly connect to F6i as depicted in the F6a case previously. This finding suggests that the nature of dynamic behaviour of fault number 6 is slightly dissimilar between the abrupt and incipient cases, despite the fact that both have a similar root of malfunction situation. In addition, the trend of identification is exactly similar regardless of number of dimensions that applied.

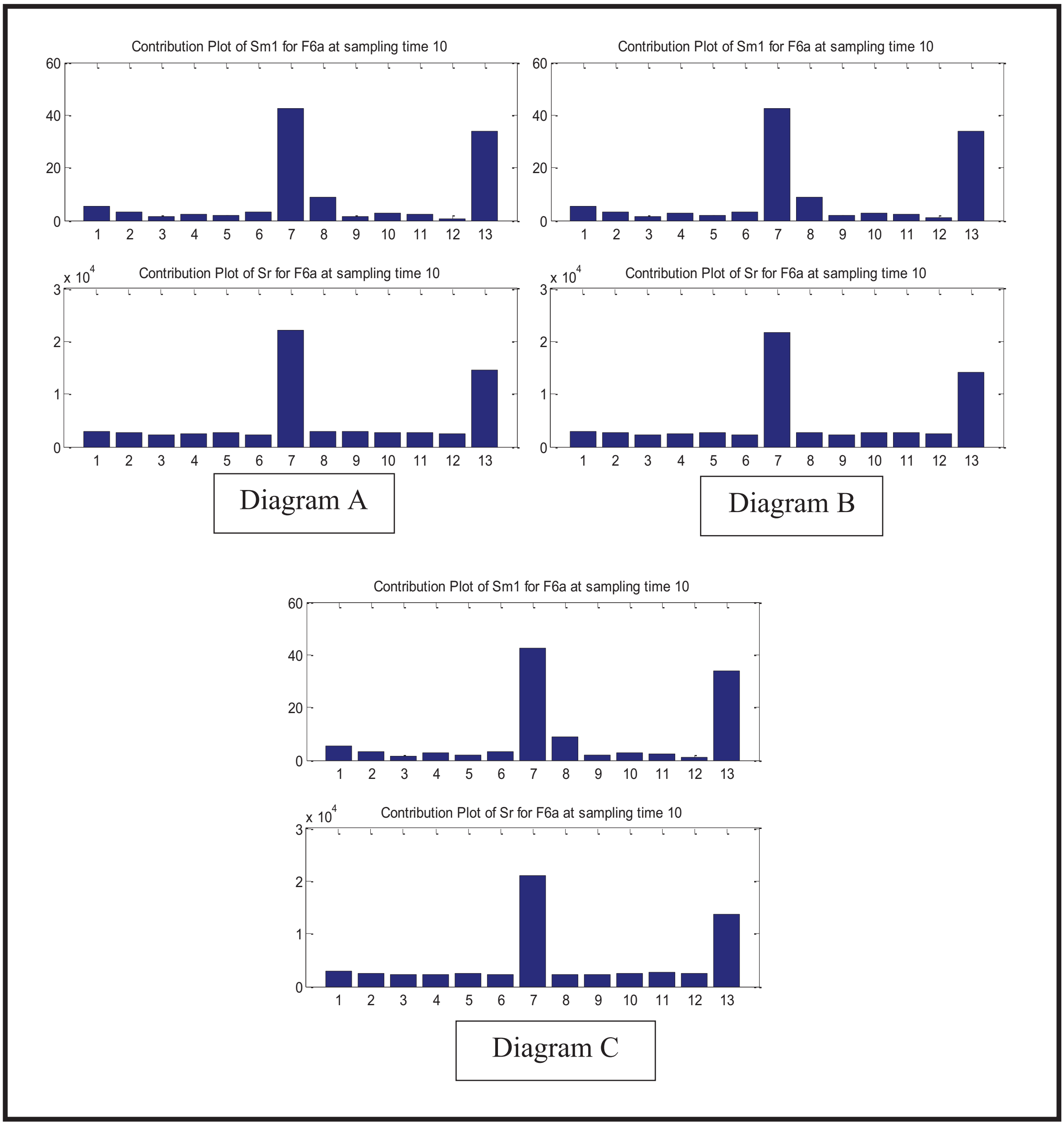

Figure 4.7: Contribution plots of S_{m1} (top) and S_r (bottom) for F6a with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C)

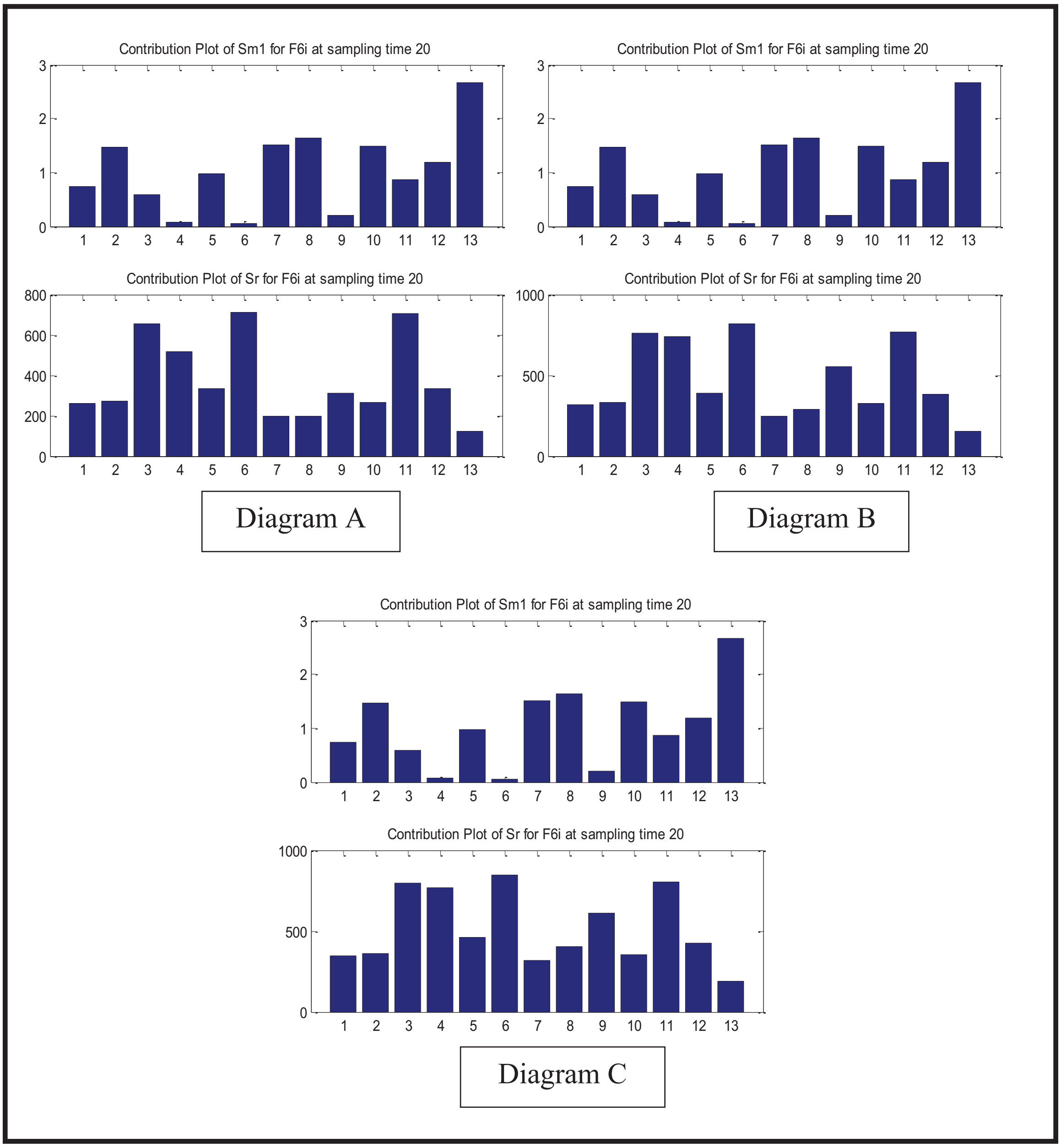

Figure 4.8: Contribution plots of S_{m1} (top) and S_r (bottom) for F6i with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C)

Figures 4.9 and 4.10 denote the monitoring statistic progressions for F9a and F9i respectively for sCMDS based monitoring system with Euclidean distance and window size 5. All the plots in Figure 4.9 show that the fault can be detected at sampling time 3. These performances are also comparable to those results of PCA (Chapter 3). Therefore, as similar to the previous case (F6a), both monitoring methods can be regarded as giving equal performance in detecting F9a. The results in Figure 4.10, however, show that sCMDS based monitoring systems give significantly better performance than PCA based monitoring system for F9i detections. As denoted in Table 4.2, the best FDT that provided by PCA is 14, whereby sCMDS managed to detect the same fault as fast as after 3 sampling time. Moreover, the results of PCA solely depend on the SPE progression, whereby T^2 has failed in all models. Nevertheless, both monitoring statistics of sCMDS are not just capable in detecting F9i, but more importantly performed at equal time of detection. This observation suggests that the basic concept of S_{m1} and S_r supports each other.

The results on fault identification of F9a as well as F9i are also determined in parallel to the outcomes of fault detection that adopted sCMDS schemes discussed previously as shown in Figures 4.11 and 4.12. All those contribution plots in Figure 4.11 have signified variables 5 and 12 as the main contributors for F9a particularly at sampling time 5 and onwards (the results of contribution plot at earlier samples have only indicated variable 5 that particularly connected to the problem, and besides, the magnitude of variable 12 has been also found gradually increased over time). In the second case, however, it is solely variable 12 which has been found importantly related to F9i, and that can be observable through sampling time 10 based on S_{m1} progression (top diagrams in Figure 4.12). This is another example to show that the dynamic trending between abrupt and incipient fault is slightly different. This particular trending has been also observable on the later sampling time, whereby the earlier progressions cannot be identified as clear as at sampling time 10. That particular pattern cannot be seen, however, on the S_r plot based on the similar period of time (bottom diagrams in Figure 4.12). In reflecting to the descriptions of the faults that presented in Chapter 3, the trends that depicted on those contribution plots of Figures 4.11 and 4.12 are explainable. This fault will lead to a reduction in the recycle flow rate (variable 5). The recycle flow controller then attempts to increase recycle flow rate by increasing the controller output (variable 12). It is also noticed that the pattern of deviation is not significantly affected by increasing the number of dimensions.

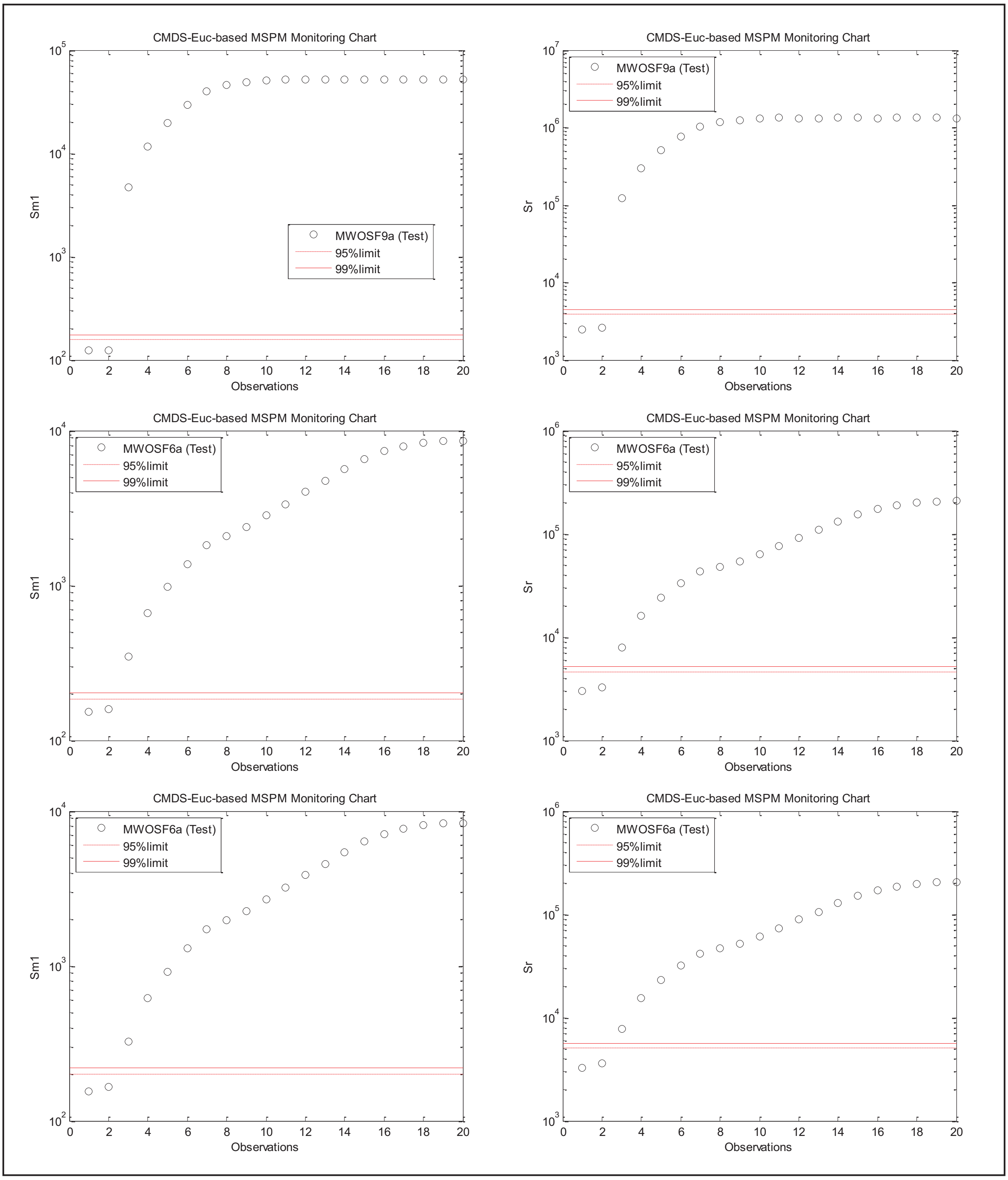

Figure 4.9: Monitoring progression of S_{m1}(left) and S_r (right) on F9a based on sCMDS using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

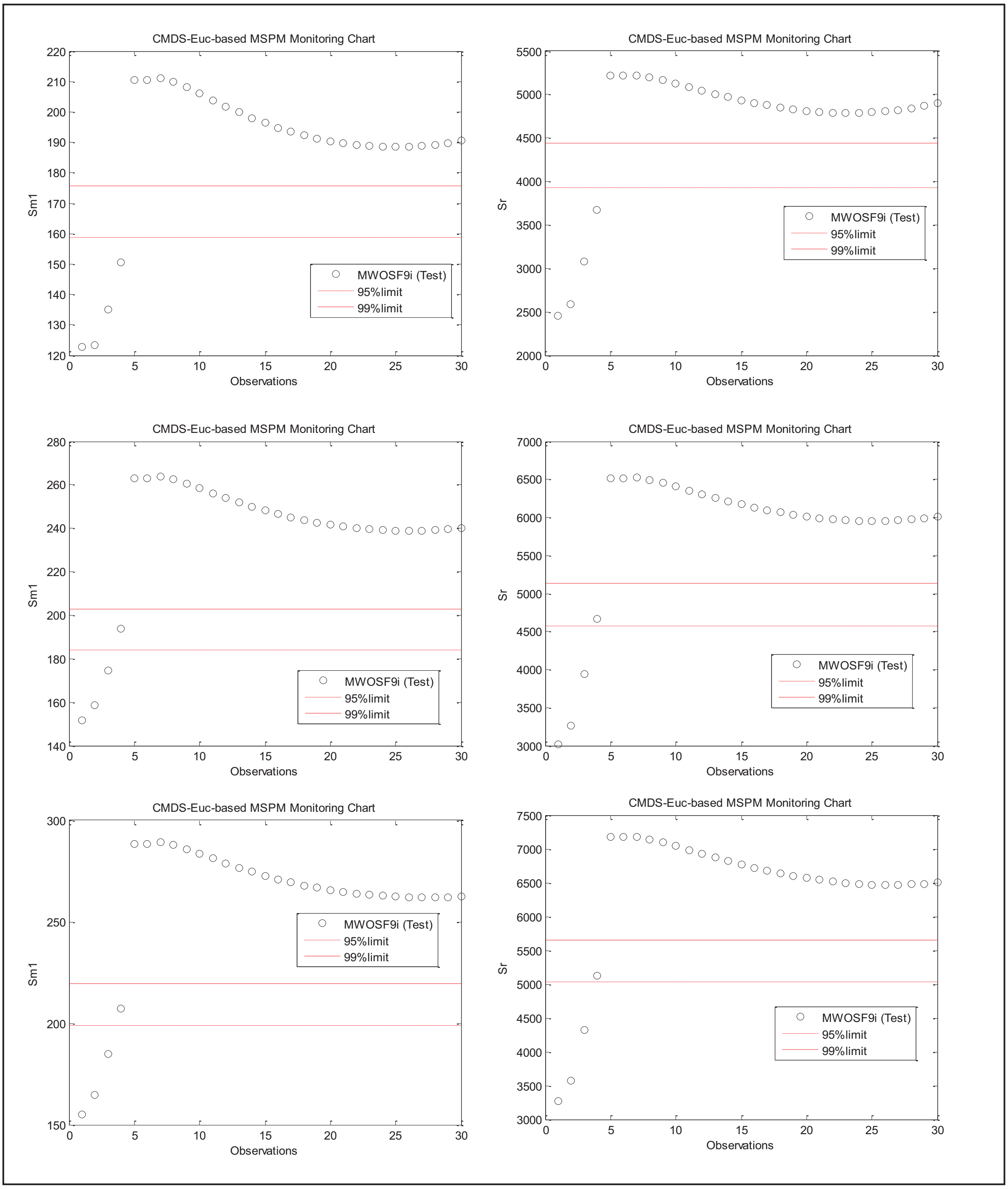

Figure 4.10: Monitoring progression of S_{m1}(left) and S_r (right) on F9i based on sCMDS using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

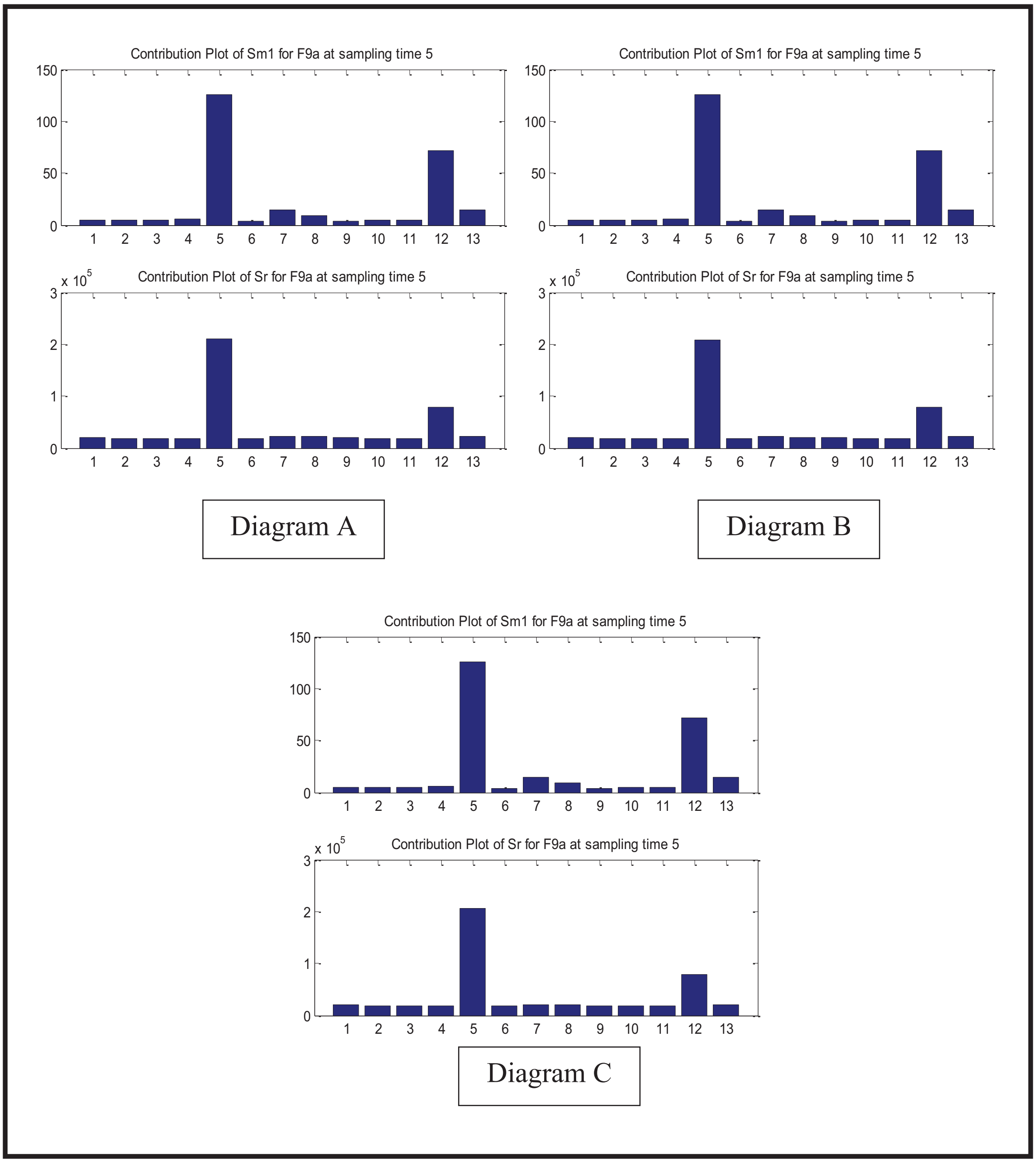

Figure 4.11: Contribution plots of S_{m1} (top) and S_r (bottom) for F9a with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C)

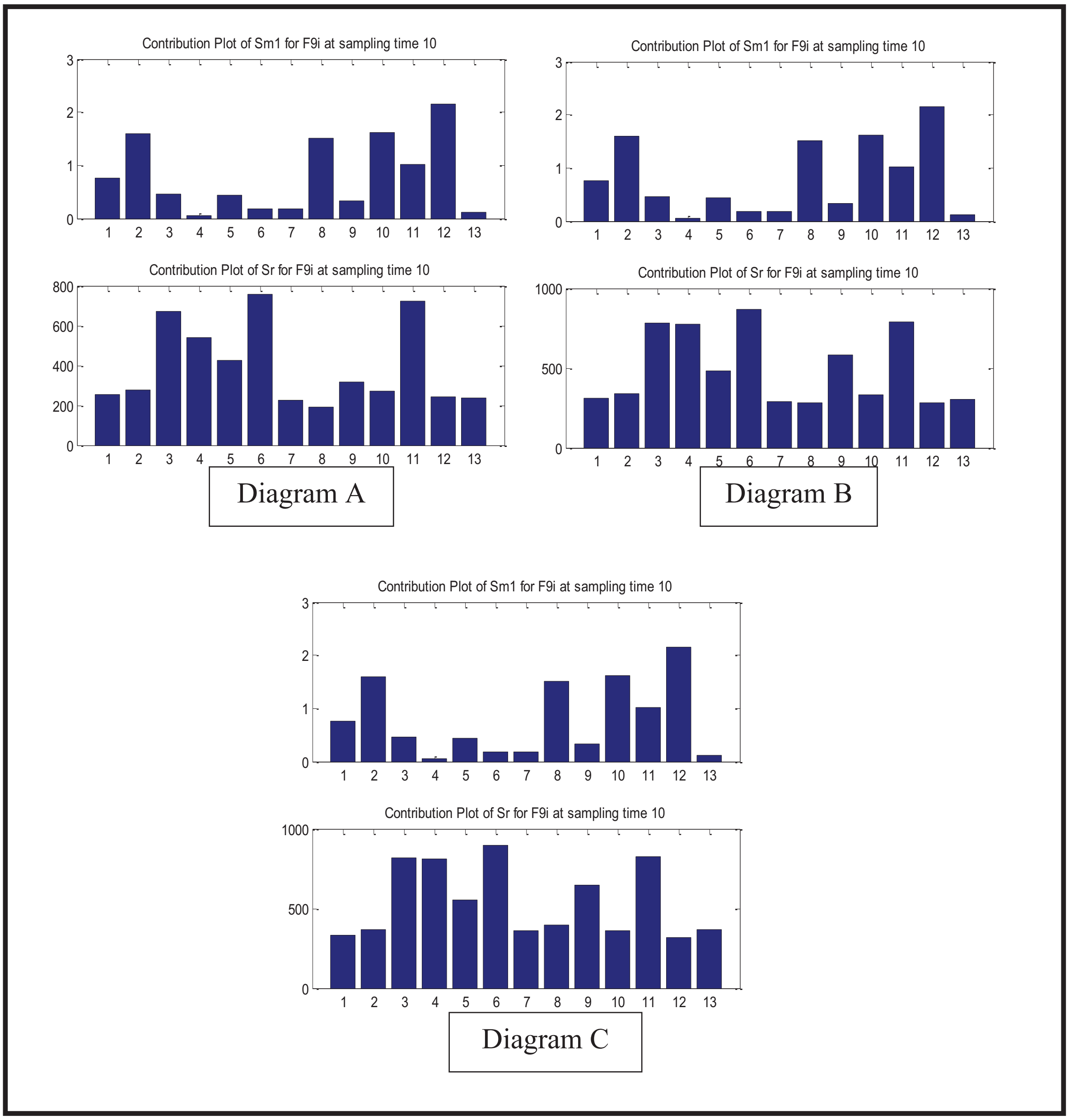

Figure 4.12: Contribution plots of S_{m1} (top) and S_r (bottom) for F9i with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C)

Lastly, Figures 4.13 and 4.14 summarize the results for F11a and F11i respectively from sCMDS based monitoring systems with Euclidean distance and window size 5. F11a can be detected by all models directly at sampling time 3, which is 1 sampling time after the fault was introduced in the process. The similar performance was also shared by PCA which has been discussed in Chapter 3. Thus, both approaches can be perceived as giving equivalent performance with regard to monitoring on this particular case. In analysing the results of F11i, however, sCMDS has again demonstrated very impressive performances. In particular, all models have managed detecting F11i as efficient as at sampling time 5 through both statistics, which is 3 sampling time delay after the fault introduced into the process. The linear PCA based monitoring system can also produce the equal fault detection performance through utilising 3 or 5 PCs. For this particular fault, both PCA and sCMDS based monitoring systems give similar performance.

The results of fault identification based on the contribution plot technique are shown in Figures 4.15 and 4.16 for abrupt and incipient fault cases respectively. All plots have denoted that variable 9 (inlet concentration) as the main contributor to the problem. These findings are determined compatible to the nature of the faults that described in Chapter 3. More importantly, the cause of the faults can be identified much earlier compared to the faults that discussed previously. In particular, the root problem of F11a and F11i can be identified as efficient as at sampling time 3 and 5 respectively (which is equal to the time of detection). All of these identifications can be performed effectively based on both statistics in connection to F11a, whereas this behaviour can be only delivered by S_{m1} with regard to F11i. Again, increasing number of dimensions does not provide any advantage as the trend for all dimensions that selected are almost identical from one to another.

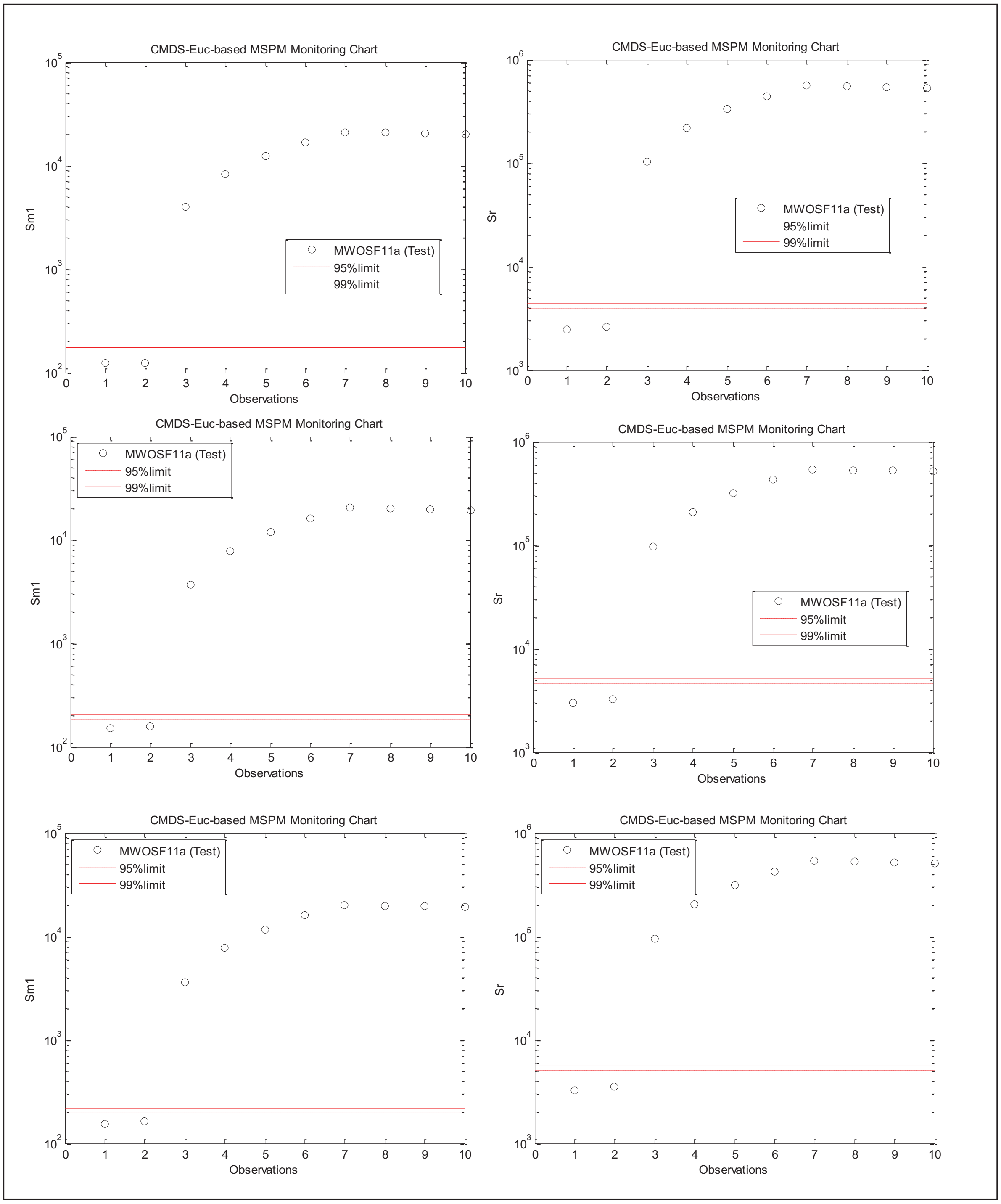

Figure 4.13: Monitoring progression of S_{m1}(left) and S_r (right) on F11a based on sCMDS using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

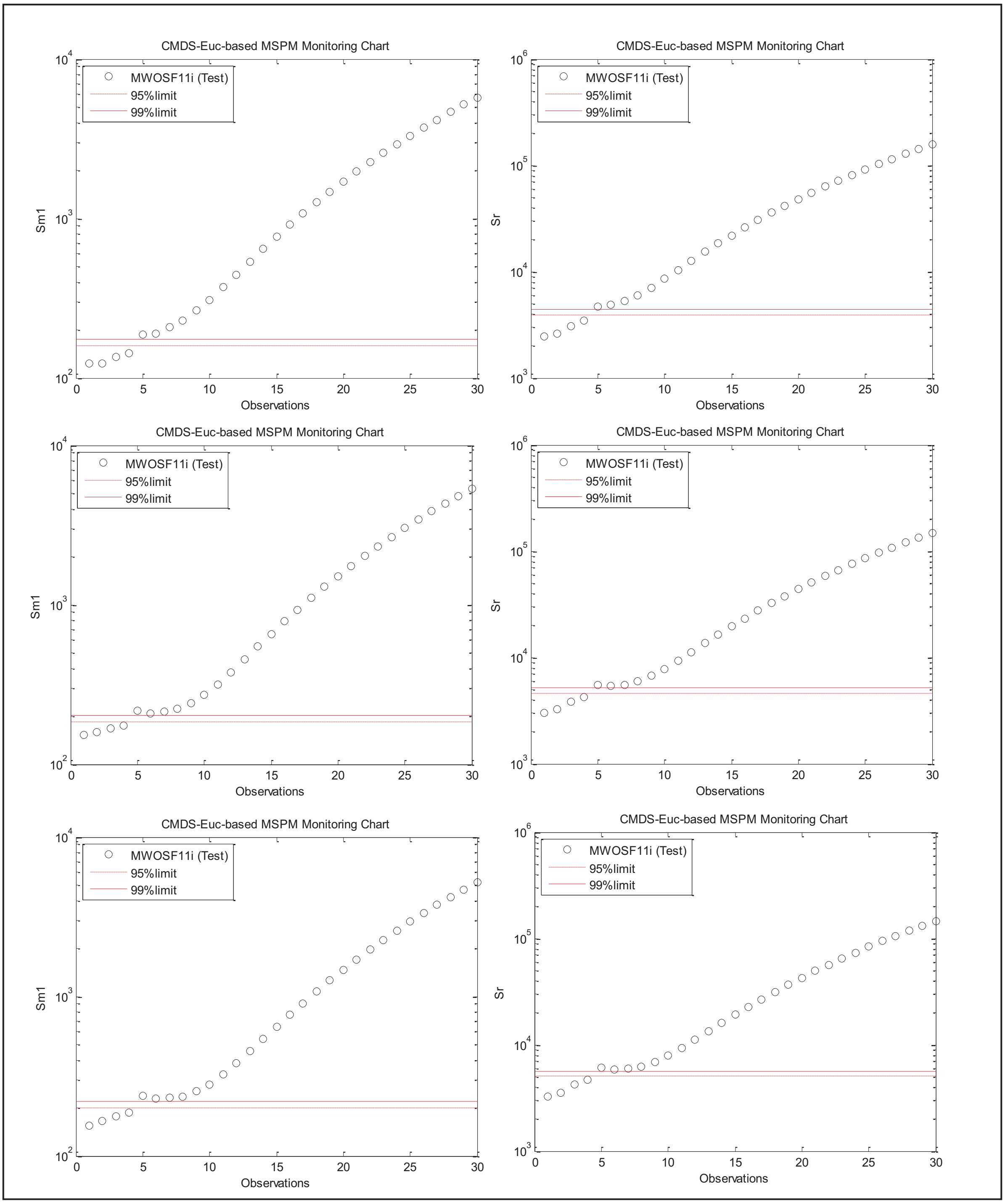

Figure 4.14: Monitoring progression of S_{m1}(left) and S_r (right) on F11i based on sCMDS using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

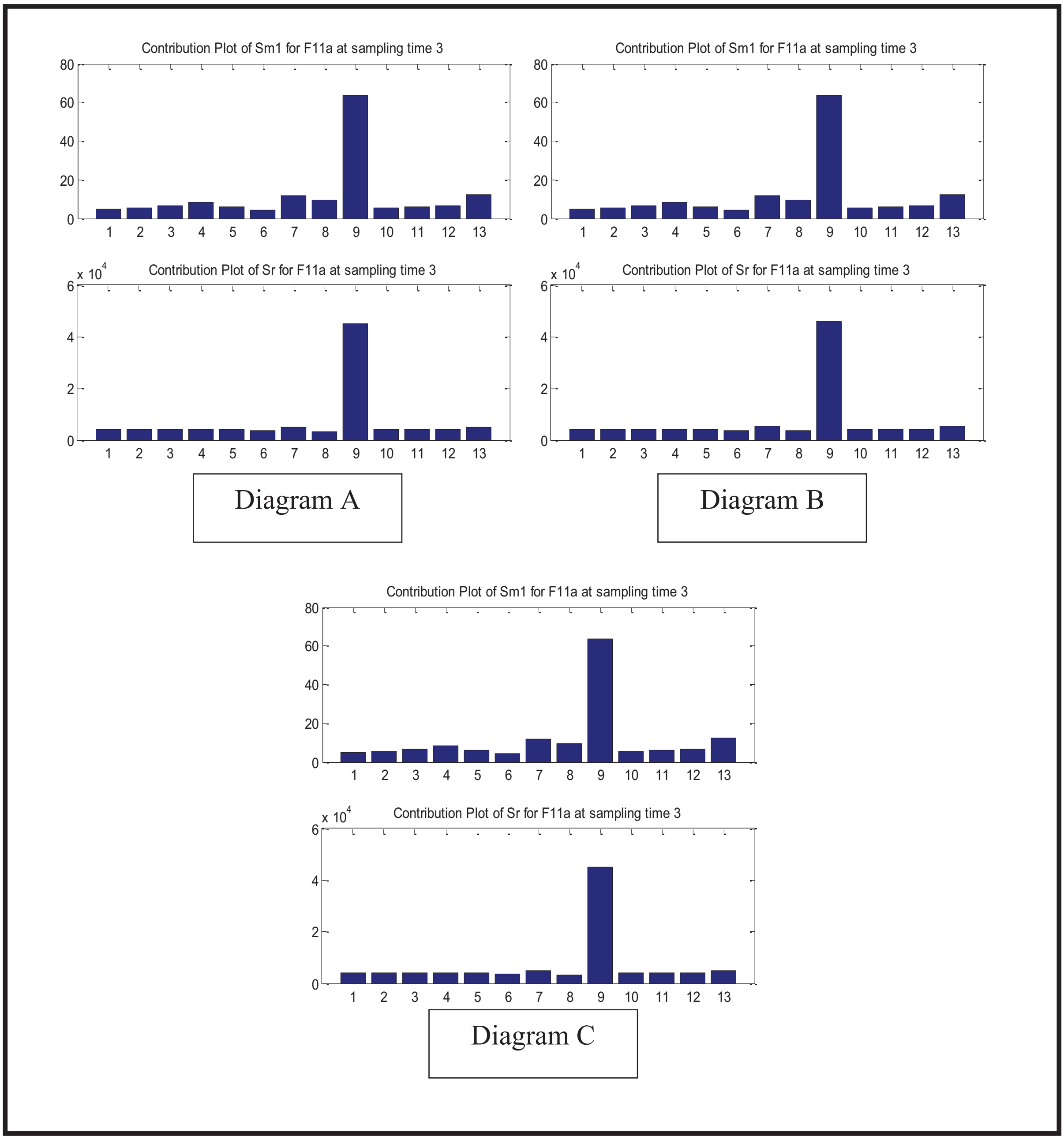

Figure 4.15: Contribution plots of S_{m1} (top) and S_r (bottom) for F11a with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C)

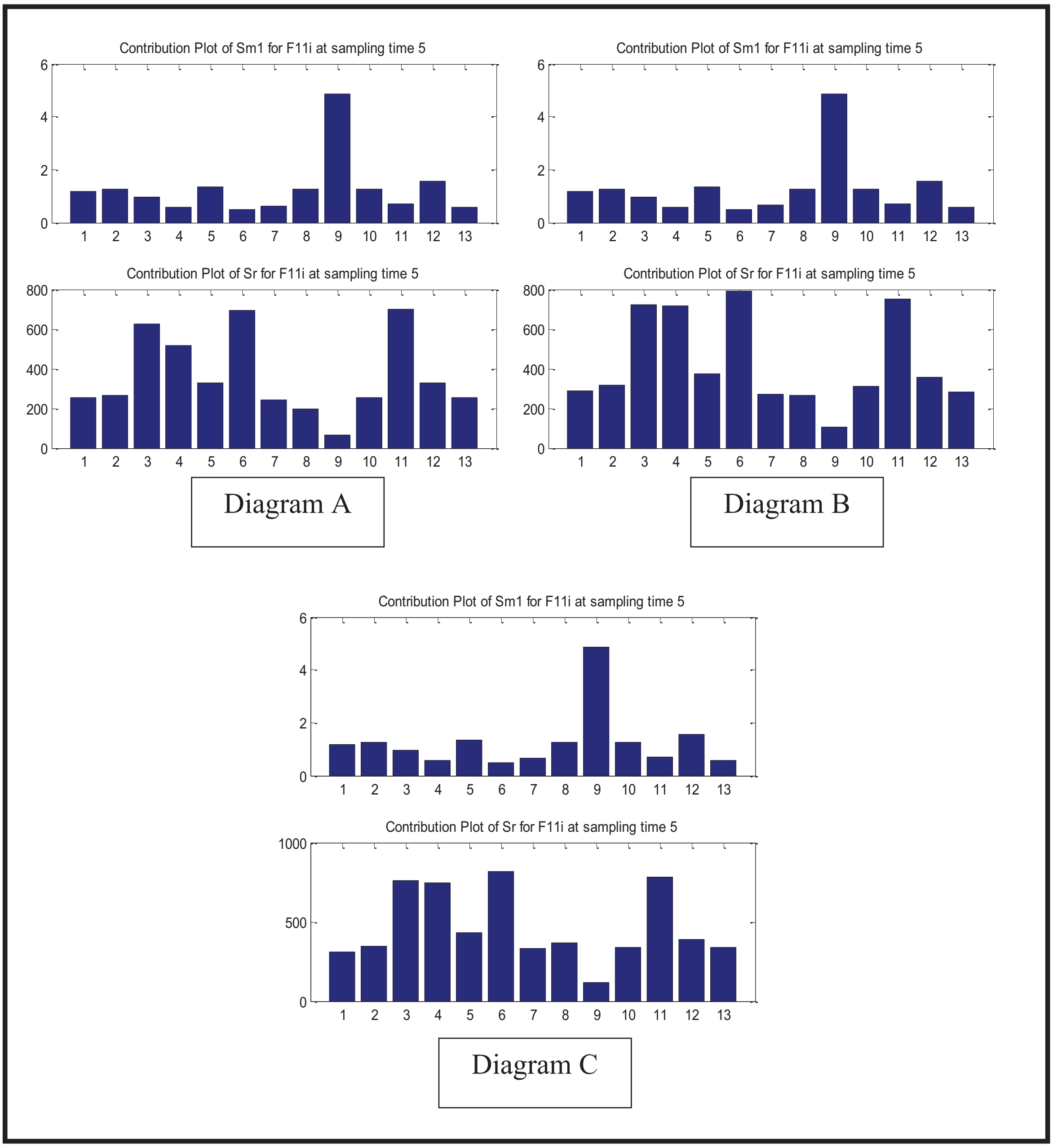

Figure 4.16: Contribution plots of S_{m1} (top) and S_r (bottom) for F11i with 3 dimensions (Diagram A), 5 dimensions (Diagram B) and 7 dimensions (Diagram C)

4.4 Results Discussion

4.4.1 The Impact of Dissimilarity Measures on The Monitoring Outcomes

This sub-section addressing the first question presented in Section 1.2 of Chapter I. Two dissimilarity measures have been used in this chapter to develop the relationships among the variables through the structure of MDS. Through analysing the results on the total number of detected cases for both distance measures, the overall trend suggests that that the performance of City-block can be considered comparable to the Euclidean scale.

The generic outcomes of the total number of fastest detection cases, on the other hand, signifies that majority of the cases show that Euclidean distance measure is superior to the City-block distance measure. In considering all of these findings, this study suggests that using the Euclidean distance measure is advantageous compared to City-block distance measure. The basic reason is that CMDS fundamentally uses the Euclidean space in developing the scores. Thus, all the scores corresponding to the City-block measure are actually representing a set of embedded Euclidean distances instead of the real City-block distance. Consequently, measuring the variables relationships' consistency in terms of City-block distance measure, between the original NOC set and the MWOS models introduce more complexity as well as lacking in its originality.

4.4.2 The Impact of using New Monitoring Statistics on The Monitoring Outcomes

This sub-section is corresponding to the second question presented in Section 1.2 of Chapter 1. Two new monitoring statistics were introduced as a result of using different approach as well as structure in developing the multivariate scores. When considering the overall performance on CSTRwR, the suggested statistics found to be efficiently as well as effectively working as expected. Nonetheless, the results were not consistent all the time, where a number of ND observed from Tables 4.1 and 4.2.

This study cannot directly or concretely provide the right answer for this situation. However, this study suspects that this could be due to the inconsistency in terms of specifying the consistent sets of eigenvectors during score development. In general, eigenvectors are subject to be rotated and reflected at the same time as a result of continuous changing of the samples by using the moving window mechanism. The impact of this situation can be seen based on a hypothetical example shown in Figure 4.17.

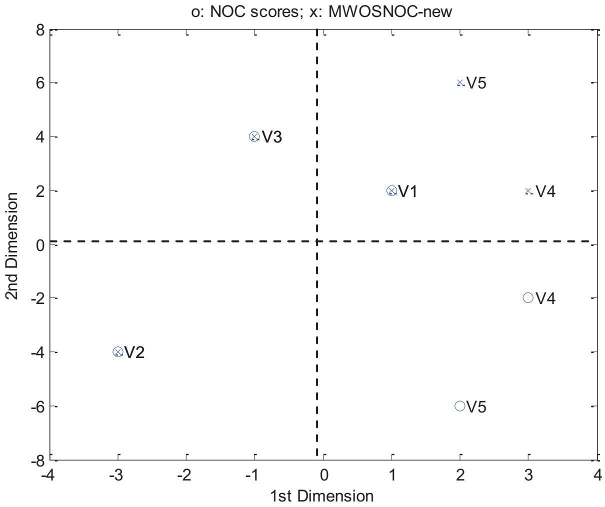

Figure 4.17: NOC scores ('o') vs MWOS-new scores ('x')

Figure 4.17 shows the new coordinates of 'V1', 'V2' and 'V3' are overlapping with the original coordinates. On the other hand, the new locations of 'V4' and 'V5' ('x' marks) have been flipped by the horizontal axis, and as a result, relocating to the opposite direction of the original configuration. This can happen as a result of sCMDS algorithm has to apply recurring different sets of eigenvector measures every time during the score projection. In other words, sCMDS may employ different settings of eigenvectors (which have been rotated and flipped), even though it involves NOC data. The negative impacts will be severed upon the second statistics (variable relationships) as the new locations of 'V4' and 'V5' will increase the S_r values. Thus, this could be the main situation that contributes to the occurrences of ND rates because the squared errors on NOC samples cannot be calculated as minimum as possible in the first phase of monitoring.

4.4.3 The Impact of Applying Various Window Settings on The Monitoring Outcomes

This sub-section is trying to deliver the answer based on the third question presented in Section 1.2 of Chapter 1. Besides of using the sCMDS algorithms, three arbitrary window size settings were applied on each of the proposed systems. The analysis was made under the assumption that reducing the window size can be regarded as the effort of decreasing the dependency on the original NOC data behaviour upon the new samples. As a consequence, the main behaviour of the MWOS samples is simply influenced by both of the sample sets and not severely dependable on the original sample set character. It was believed that, if such situation can be produced, a random behaviour of NOC data can be achieved and, eventually, the monitoring statistics comply to the chi-squared distribution assumption and more precise control limits can be produced. At the same time, it is also believed that this approach may help the system to be more sensitive during fault detection operation. From the results, the FAR outputs on the CSTRwR cases have been shown to be zero. The overall results of total number of detected cases show that the impact on using various window settings was not that obvious because those three window sizes used proved in detecting all cases (subject to the type of scales applied). However, using smaller window size may help the monitoring system to gain faster detection (total number of detection cases) especially on those cases that slowly detected by PCA (incipient faults).

4.4.4The Impact of Applying Smaller Dimensionalities on The Monitoring Outcomes

This sub-section is pertaining to the last question presented in Section 1.2 of Chapter 1. This is the key to the particular motive of this study. In particular, the linear-PCA was criticized for being as problematic as the technique involves a high number of PCs for effectively modelling non-linear processes. In other words, the conventional method requires relatively large number of PCs for capturing the variations of the original process data and eventually it can detect the faults in faster pace as well as in a more productive manner compared to the performance of monitoring systems with smaller number of PCs. These assumptions were verified based on the results observed on the CSTRwR cases.

However, the generic performance of the sCMDS system generally suggests that there is no significant improvement in terms of the fault detection outcomes as a result of applying more dimensionalities into the monitoring model. In fact, there were cases still undetected by using higher dimensions (Tables 4.1 and 4.2). Although this can be perceived as a kind of improvement to some extents, but the inconsistency is permanent and cannot be removed by simply adopting more dimensionalities.

Unlike to the conventional approach, CMDS uses the variable themselves as the main score objects in the reduced dimensional space. At the same time, the inter-distance measures are also utilised to represent the conceptual association among the variables. Therefore, the only information that is transferred from the original data space to the latent space is on the structure of variable configuration instead of sample variances (corresponds to the traditional concept). Thus, increasing the number of dimensions will only contribute in enhancing the variable correlations (numerically) relative to the original and not on the nature of the sample distribution. The main benefit of this approach is that the variable correlations can be designed at any scales of dimensions, but preferably in a lesser quantity. This has been proven in the case of CSTRwR, where all of the specified cases can be detected by all the Euclidean-based sCMDS systems using dimension 3. Besides, the incipient faults can still be detected relatively faster than PCA using only three dimensions. Nonetheless, the optimum amount of sufficient number of dimensions which can be configured with the number of variables and window size applied is still unknown. Thus, the current method of specifying the number of dimensions which is based on the ratio of eigenvalues will be maintained.

4.5 Summary

This chapter presents the results of using sCMDS for monitoring with respect to different sets of settings demonstrated on the CSTRwR system. The outcomes were also compared quantitatively with the performances from the linear PCA based monitoring system. It was shown that sCMDS has proven working comparative to the PCA based monitoring system performance particularly based on abrupt fault cases. However, sCMDS has been shown to give faster detection than the monitoring performance of linear PCA model based monitoring systems with regard to incipient fault cases. Nonetheless, there were some minor inconsistencies observed based on the sCMDS results. Thus, the following chapter presents another type of CMDS-based framework, which is an enhanced strategy compared to the current method.

CHAPTER 5

FRAMEWORK II: MDS-BASED MSPM SYSTEM USING MOVING WINDOW CMDS PROCRUSTES ANALYSIS PROJECTION

5.1 Introduction

This chapter presents the methodology as well as monitoring results of framework II, Classical Multidimensional Scaling-Procrustes Analysis-based (CMDS-PA) MSPM system. The results of framework I discussed in Chapter 4 suggested that an improvement should be made on the sCMDS algorithms in order to obtain higher degree of consistency in the fault detection outcomes. It was assumed that those inconsistencies are generated from the variation in terms of eigenvector measures, which eventually may modify the correct reconfiguration of MWOS-NOC scores as well as the corresponding limits that developed. Thus, an enhanced technique, which focuses on standardized PA transformation factors, is proposed in order to improve the previous CMDS monitoring framework results. The background idea of this approach is to emulate the approach of loading factors obtained from the PCA algorithms.

One of the main techniques available in the domain of MDS that addresses specifically the score reconfiguring procedures is known as Procrustes Analysis (PA). According to Cox and Cox (1994), PA can be regarded as one-to-one mapping technique that matches from a set of configurations to another by producing a measure of fitness. Cox (2001) has utilized PA for matching two configuration sets (one was produced by CMDS, while another has been developed through Non-metric MDS) for monitoring, and the results show that the mapped configuration closely matched the targeted set. Therefore, the aim of the second framework is to utilize PA transformation functions and CMDS to develop the scores as well as evaluating the impacts of the approach on the fault detection performance.

This chapter is organized as follows. Section 5.2 explains the detail procedures of the proposed methodology while Section 5.3 presents the monitoring results on the CSTRwR, taking into account the comparative performance between PCA and the first framework. Finally, a summary is presented to summarize the key findings.

5.2 Methodology

The detail procedures of this framework are illustrated in Figure 5.1.

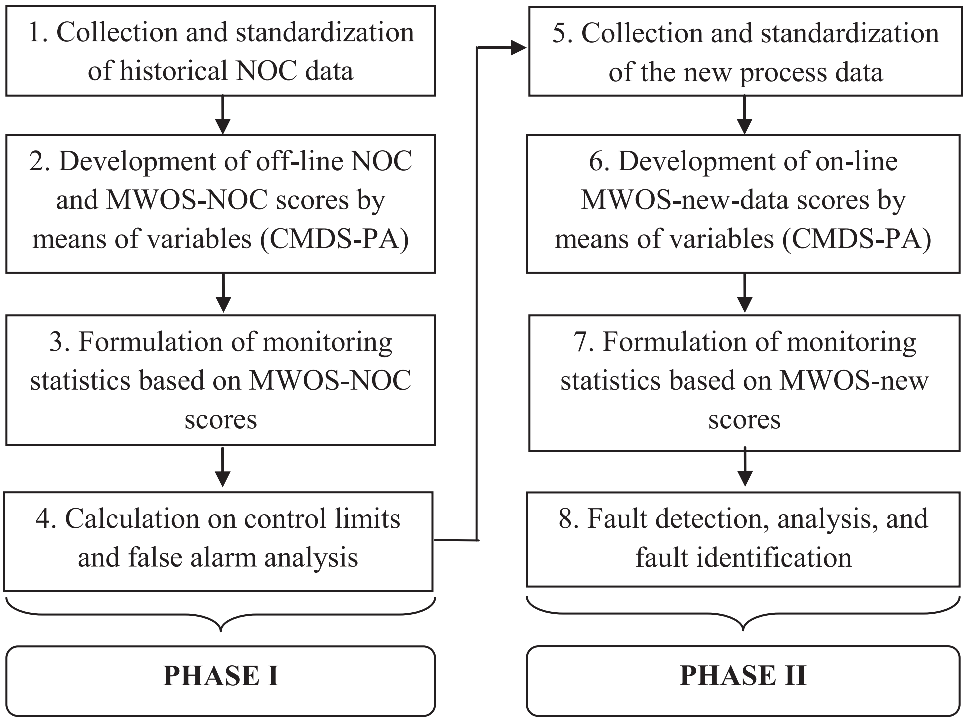

Figure 5.1. CMDS-PA-based MSPM framework

Compared to framework I (Figure 4.1), the main differences are on steps 2, 3, 6 and 7, where the projection of the scores will be executed through applying CMDS and PA procedures at the same time.

5.2.1 Phase I Procedures

After executing the first step, the second step basically involves multivariate score development. Firstly, the standard CMDS algorithms are applied to both sets of the original NOC scores, $\mathbf{X}_{\mathbf{NOC1}}$ and $\mathbf{X}_{\mathbf{NOC2}}$, to produce multivariate scores for the first set, $\mathbf{Y}_{\mathbf{NOC1}}$ and multivariate scores for the second set, $\mathbf{Y}_{\mathbf{NOC2}}$ respectively. Then, the PA procedures are applied as described in the following (Borg and Groenen, 1997):

i. Computation of the minor product moment between the first NOC scores, $\mathbf{Y_{NOC1}}$ (originated from $\mathbf{X_{NOC1}}$), and the second NOC scores, $\mathbf{Y_{NOC2}}$ (originated from $\mathbf{X_{NOC2}}$).

$$\mathbf{C_{PA}}=\mathbf{Y_{NOC1}}^{T}\mathbf{J_m}\mathbf{Y_{NOC2}} \tag{5.1}$$

where $\mathbf{J_m}$ is given in Equation 2.41.

ii. Decomposition of $\mathbf{C_{PA}}$ into the eigen basic structures.

$$\mathbf{C_{PA}}=\mathbf{P_{PA}}\mathbf{V_{PA}}\mathbf{P_{PA}}^{T} \tag{5.2}$$

where $\mathbf{P_{PA}}$ and $\mathbf{V_{PA}}$ are the eigenvectors and eigenvalues matrices of $\mathbf{C_{PA}}$ respectively.

iii. Identification of the optimal rotation matrix:

$$\mathbf{R} = \mathbf{P_{PA}}\ \mathbf{P_{PA}}^{T} \tag{5.3}$$

iv. Identification of the optimal dilation scale:

$$s=(\mathrm{tr}\mathbf{Y_{NOC1}}^{T}\mathbf{J_m}\ \mathbf{Y_{NOC2}}\mathbf{R})/(\mathrm{tr}\ \mathbf{Y_{NOC2}}^{T}\mathbf{J_m}\mathbf{Y_{NOC2}}) \tag{5.4}$$

v. Identification of the optimal translation vector:

$$\mathbf{t}=(\mathbf{Y_{NOC1}}-s\ \mathbf{Y_{NOC2}}\mathbf{R})^{T}\mathbf{1}/m,\ m=\text{number of variables} \tag{5.5}$$

vi. The reproduction of MWOS-NOC scores projected by PA for sample '*k*' is given by:

$$\mathbf{Y_{PA\text{-}NOC(\mathit{k})}}=s\ \mathbf{Y_{MWOS\text{-}NOC(\mathit{k})}}\mathbf{R}+\mathbf{1}\mathbf{t}^{T} \tag{5.6}$$

Where $\mathbf{Y_{MWOS\text{-}NOC(k)}}$ = new scores of MWOS-NOC samples at time *k*.

These particular CMDS-PA procedures are slightly different compared to the algorithms proposed in Yunus and Zhang (2010b; 2010c). The procedures are described as follow:

i. Computation of the minor product moment between the reconstructed NOC matrix, $\mathbf{Y}$ and the modified NOC matrix, $\mathbf{X_{mod}}$: $\mathbf{C_{PA}}=\mathbf{Y}^{T}\mathbf{J_m}\mathbf{X_{mod}}$ (5.7)

where $\mathbf{J_m}$ is from Equation 2.41 and $\mathbf{X_{mod}}$ is a modified NOC data with size *m* by *p*.

ii. Application of the eigen decomposition on $\mathbf{C_{PA}}$ by way of $\mathbf{C_{PA}}=\mathbf{P_{PA}}\mathbf{V_{PA}}\mathbf{P_{PA}}^{T}$ (5.8)

where $\mathbf{P_{PA}}$ and $\mathbf{V_{PA}}$ are respectively eigenvector and eigenvalues matrices of $\mathbf{C_{PA}}$.

iii. Calculation of the optimal rotation matrix, $\mathbf{R} = \mathbf{P_{PA}} \mathbf{P_{PA}}^T$. (5.9)

iv. Calculation of the optimal dilation scale, s=(tr$\mathbf{Y}^T\mathbf{J}\mathbf{X_{mod}}$ $\mathbf{R}$)(tr$\mathbf{X_{mod}}^T\mathbf{J}\mathbf{X_{mod}}$). (5.10)

v. Calculation of the optimal translation vector, $\mathbf{t}=(\mathbf{Y}-s\mathbf{X_{mod}}\mathbf{R})^T\mathbf{1}/m$. (5.11)

vi. The final transformed model of NOC is given by

$$\mathbf{Y_{PA(k)}}=s\mathbf{X_{MWOS\text{-}mod(k)}}\mathbf{R}+\mathbf{1t}^T \quad (5.12)$$

Where, $\mathbf{X_{MWOS\text{-}mod(k)}}$ = MWOS-NOC samples at time k.

As provided in Equations 5.7 to 5.12, both of these previous works were focusing on transformation between one set of NOC scores in the reduced dimensional space (**Y**) and series of the original MWOS-NOC data which is structurally modified by the use of average function ($\mathbf{X_{mod}}$). The purpose of using the averages is to smooth out measurement noise. Thus, the main drawback of this former approach can be spotted by the changes in the original data behaviour, where the data were forced by forming a line configuration instead of a cluster (based on a plot in two dimensional bases). In contrast, the one described in this chapter takes into the consideration the transformation between two sets of NOC scores in the compressed dimensional space, where no structural adjustment is imposed initially. Hence, the true data behaviour of both data sets is preserved accordingly by forming clustering objects which are relevant to the assumption made.

The next step concerns with the monitoring statistic formulation. In particular, framework II utilizes the second type of the first statistic (S_{m2}) which is shown in Equation 5.13 representing the sum of squared errors in terms of projected variables in the reduced dimensional space between $\mathbf{Y_{PA}}$ and $\mathbf{Y_{PA\text{-}NOC1}}$ configurations.

Statistic 2 for means (S_{m2}): Sum of squared errors in terms of projected variables in the reduced dimensional space.

$$S_{m2} = \sum_{i=1}^{m}\left(\sum_{j=1}^{p}\left[y_{PA}(i,j) - y_{PA-NOC1}(i,j)\right]^2\right) \quad (5.13)$$

The use of S_{m2} instead of S_{m1} is imperative because it suits to the main objective of PA implementation. More specifically, PA tries to match between two sets of configurations such that the locations of the second set can be modified according to the coordinates of the targeted set. In contra to this approach, the scores in the first framework are developed totally depending on the mapping between two different sets of dissimilarity (distances) measures instead of coordinates. Thus, the only effective way to measure the scale of fitness by using the PA approach is by analyzing the magnitude of S_{m2}, which is graphically shown in Figure 5.2.

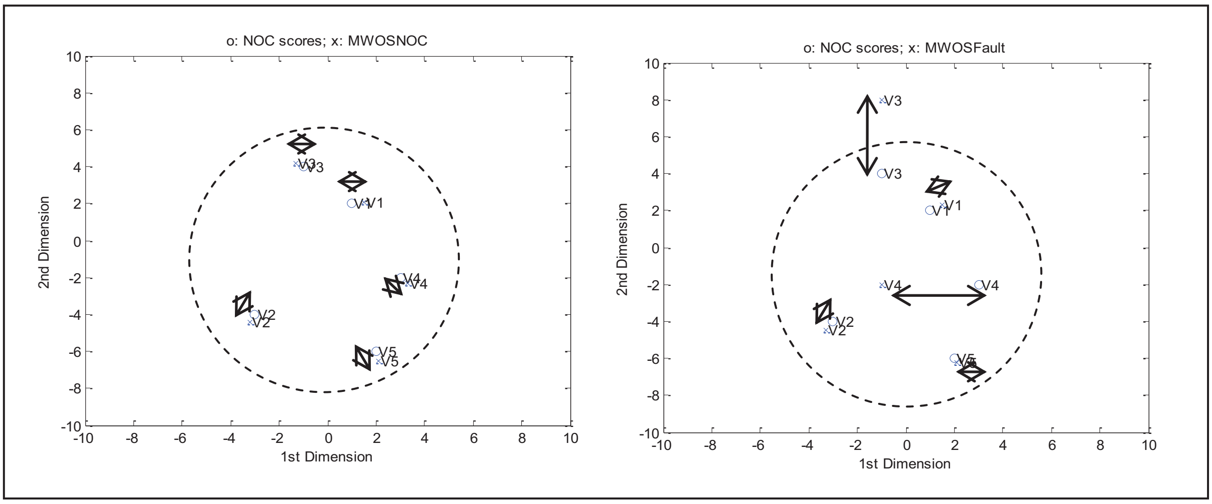

Figure 5.2: Illustration of S_{m2} based on the plots of NOC scores vs MWOS-NOC scores (left diagram); NOC scores vs MWOS-fault scores (right diagram)

The left graph of Figure 5.2 depicts that all the MWOS scores are relatively close in terms of coordination compared to the original NOC scores. The reason is basically similar to the previous discussion in Chapter 4. As for the right diagram of Figure 5.2, on the other hand, the runaway impacts that are represented by 'V3' and 'V4' can be easily highlighted by the use of S_{m2}. In particular, S_{m2} will be affected by having a large magnitude corresponding to both cases. Therefore, S_{m2} has a significant advantage as opposed to S_{m1}, particularly concerning on the 'V4' faulty condition. Anyway, both S_{m1} and S_{m2} share the same basic information, where they relate to the concept of magnitude of deviation of the current sample (represented by a group of samples through MWOS mechanism) from the pre-defined points (centre) of NOC scores.

The same statistic S_r from Equation 4.2 is utilized again in this framework. Basically, this parameter measures the sum of squared errors in terms of dissimilarity measures between the $\mathbf{Y_{PA\text{-}NOC}}$ and $\mathbf{Y_{NOC1}}$ configurations, which conceptually analysing the consistency of variables correlations. The particular S_r version applied for framework II is shown in Equation (5.14).

$$S_r = \sum\left(\sum_{i=1}^{m}\sum_{j=1}^{m}\left[d_{PA(ij)} - \delta_{PA-NOC1(ij)}\right]^2\right) \; ; i \neq j \tag{5.14}$$

Besides, the monitoring limits are also developed by using equation 4.3 and that is based on the chi-squared distribution. The false alarm analysis (FAR) analysis will be also conducted for evaluating the statistic robustness.

5.2.2 Phase II Procedures

The phase II procedures basically follow the same order as in the first framework in Chapter 4. Anyway, the projecting algorithms for the new sample scores employed by CMDS-PA have been improved from sCMDS procedures, particularly through adapting the PA technique as shown in Equation 5.15.

$$\mathbf{Y}_{\mathbf{PA\text{-}new}(k)} = s\mathbf{Y}_{\mathbf{MWOS\text{-}new}(k)}\mathbf{R} + \mathbf{1}\mathbf{t}^{T} \tag{5.15}$$

where, $\mathbf{Y}_{\mathbf{PA\text{-}new}(k)}$ = reproduction of the scores projected by PA for MWOS-new sample 'k'.

$\mathbf{Y}_{\mathbf{MWOS\text{-}new}(k)}$ = multivariate scores of MWOS sample 'k' developed by CMDS.

The monitoring statistics S_{m2} (Equation 5.13) and S_r (Equation 5.14) are used. The fault detection outcomes were also assessed using the previous performance indicators – total number of detected cases and total number of fastest detection cases.

Two approaches of contribution plot techniques are proposed in this framework in order to gain the generic insight of the probable variables that might contribute to the detected signal. The first applies similar conception as in framework I as indicated in Equation 5.16.

$$\left(CMDS - PA\ Statistics\right)_j = \sum_{i=1}^{m}\left(X_i\right)_j \tag{5.16}$$

Where, $(CMDS\text{-}PA\ Statistics)_j$ = CMDS statistics (S_{m2} or S_r) at a particular sampling time 'j'.

$(X_i)_j$ = contribution of the ith variable to CMDS-PA statistics at a particular sampling time *'j'*.

Equation 5.16 is applied to both statistics that proposed. On the other hand, another approach is based on differential contribution as denoted in Equation (5.17).

$$(\mathbf{dc})_j = \left[\left(Contribution_{CMDS-PA(fault)} \right)_i - \left(Contribution_{CMDS-PA(normal)} \right)_i \right]_j \quad (5.17)$$

Equation 5.17 basically shows the difference in terms of monitoring statistic contributions between the abnormal and normal operation samples of CMDS-PA statistics. In particular, **dc** (differential contribution) is a vector that contains *'m'* number of errors corresponding to the monitored process variables. Meanwhile, *'i'* is an index for process variables, whereas *'j'* pertains to sampling time. The term *'*$Contribution_{CMDS\text{-}PA(fault)}$*'* refers to S_{m2} or S_r individually, under which, the current operation contains faulty condition, whereas *'*$Contribution_{CMDS\text{-}PA(new)}$*'*, on the other hand, relates to any of NOC statistics of CMDS-PA. The main intention of employing the differential contribution plot is to gain faster identification time against the conventional technique. In other words, it is believed that that those variables which indicate large magnitude in **dc** value may have strong relationships with the detected signal, either contributing or being affected by the malfunction condition.

Both of these techniques of fault identification are executed once the detection signal is initiated by the proposed CMDS-PA monitoring system. The analysis will be conducted from the first sample of detection to a point, where the trend of deviation indicated by large magnitude bar(s) is (are) consistent over the time. All of these steps were implemented in the CSTRwR case study that discussed previously in Chapter 3.

5.3 Results and Analysis

The focus of this section is to conduct a generic overview on the quantitative comparison of fault detection performances between the current framework and PCA based monitoring system as well as taking into consideration the performances from framework I. All the malfunction cases used in Chapter 4 are utilised again to demonstrate the capability of framework II. The results are then summarized, whereby the reflections on the outcomes are provided at the end.

As explained in the previous chapter, the original NOC data was divided into two parts (each contained 50 samples), $\mathbf{X}_{\mathbf{NOC1}}$ and $\mathbf{X}_{\mathbf{NOC2}}$, where the first has been used to develop the NOC score model ($\mathbf{Y}_{\mathbf{NOC1}}$) by using the standard CMDS algorithms. It has been explained in the previous chapter that CMDS models with 3, 5 and 7 dimensions were developed to represent around 75%, 90% and 99% proximity to the original dissimilarity measures. Then, the standard CMDS procedures were again used to construct the scores for $\mathbf{X}_{\mathbf{NOC2}}$, which eventually produces $\mathbf{Y}_{\mathbf{NOC2}}$. Once this was completed, the PA procedures were applied in order to identify those transformation factors $\mathbf{R}$, s and t between $\mathbf{Y}_{\mathbf{NOC1}}$ and $\mathbf{Y}_{\mathbf{NOC2}}$. The moving window mechanism was then applied integrating both $\mathbf{X}_{\mathbf{NOC1}}$ and $\mathbf{X}_{\mathbf{NOC2}}$. Three window sizes, 5, 10 and 15 were selected to introduce random variations in the NOC sampling distribution. Then, MWOS-NOC scores were obtained through the standard CMDS procedures. Later, Equation 5.6 is applied for PA projection. Through applying equations 5.13 and 5.14, series of monitoring statistics were then established that extracted from the reproduction of the scores based on the PA technique.

Eventually, a set of control limits were then computed and a new set of NOC data (testing NOC set, $\mathbf{X}_{\mathbf{NOC\text{-}Test}}$ with size of 50 samples by 13 variables) was then used to assess the robustness of the control limits. The overall results show that FAR values on the testing NOC data are zero with regard to all model settings. This indicates that the established monitoring limits should be robust in the monitoring operation as similar to the previous framework.

Tables 5.1 and 5.2, present the overall monitoring performances in terms of fault detection time (FDT) of both CMDS-PA and PCA systems based on 3, 5 and 7 dimensions/PCs settings corresponding to abrupt and incipient fault cases respectively. In all of the cases, the faults were detected such that at least 3 successive samples of either monitoring statistics were located above their 99% control limits. Both tables have also highlighted the most effective as well as fastest detection time in bold for every case that analysed.

Table 5.1: Fault detection times of monitoring systems based on CMDS-PA and PCA for abrupt fault cases

FAULT DETECTION TIME (samples)																						
Fault Cases	Dimensions	3							5							7						
	Methods	CMDS-PA						PCA	CMDS-PA						PCA	CMDS-PA						PCA
	Window sizes	5		10		15			5		10		15			5		10		15		
	Scales	E	C	E	C	E	C		E	C	E	C	E	C		E	C	E	C	E	C	
1a	Fault Detection Time	**1**	**1**	**1**	2	**1**	2	**1**	**1**	**1**	**1**	2	**1**	2	**1**	**1**	**1**	**1**	2	**1**	2	**1**
2a		**1**	2	**1**	2	**1**	3	**1**	**1**	2	**1**	2	**1**	3	**1**	**1**	2	**1**	2	**1**	3	**1**
3a		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
4a		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
5a		**1**	**1**	**1**	2	**1**	**1**	**1**	**1**	**1**	**1**	2	**1**	**1**	**1**	**1**	**1**	**1**	2	**1**	**1**	**1**
6a		**1**	2	**1**	**1**	**1**	**1**	**1**	**1**	2	**1**	2	**1**	**1**	**1**	**1**	2	**1**	2	**1**	**1**	**1**
7a		8	8	9	10	10	12	**5**	8	8	8	10	10	12	**2**	9	8	8	12	10	12	**1**
8a		**1**	2	**1**	3	**1**	3	**1**	**1**	2	**1**	3	**1**	3	**1**	**1**	2	**1**	3	**1**	3	**1**
9a		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
10a		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
11a		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Total number of detected cases		**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**
Total number of fastest detected cases		10	7	10	6	10	7	**11**	10	7	10	5	10	7	**11**	10	7	10	5	10	7	**11**
Legends : E=Euclidean C=City-block																						

Table 5.2: Fault detection times of monitoring systems based on CMDS-PA and PCA for incipient fault cases

FAULT DETECTION TIME (samples)																						
Fault Cases	Dimensions	3							5							7						
	Methods	CMDS-PA						PCA	CMDS-PA						PCA	CMDS-PA						PCA
	Window sizes	5		10		15			5		10		15			5		10		15		
	Scales	E	C	E	C	E	C		E	C	E	C	E	C		E	C	E	C	E	C	
1i	Fault Detection Time	**2**	3	3	3	3	4	**2**	**2**	4	3	4	3	5	4	**2**	3	3	3	3	5	5
2i		4	5	6	6	6	8	**3**	**3**	5	6	6	6	8	7	**3**	5	6	6	7	8	8
3i		3	5	4	6	4	6	**2**	3	5	3	6	4	5	**2**	3	5	3	6	4	6	**1**
4i		**3**	ND	26	28	31	29	29	**2**	3	8	7	28	26	25	**2**	3	7	6	27	12	15
5i		**3**	13	15	16	18	19	10	**3**	13	8	8	17	18	9	**3**	**3**	8	8	17	14	22
6i		**3**	44	26	8	36	35	40	**2**	3	8	7	32	30	34	**2**	3	7	6	31	12	21
7i		**3**	41	26	8	36	34	39	**2**	3	8	6	32	30	32	**2**	3	7	6	33	12	21
8i		**3**	6	24	8	37	36	27	**2**	3	8	6	34	33	24	**2**	3	7	6	33	12	15
9i		**3**	43	18	8	29	28	29	**2**	3	8	7	26	23	24	**2**	3	7	6	24	12	14
10i		**3**	38	26	8	36	34	30	**2**	3	8	6	33	31	25	**2**	3	7	6	31	12	16
11i		**3**	5	5	6	6	8	**3**	**3**	6	5	7	6	8	**3**	**3**	6	5	7	6	8	9
Total number of detected cases		**11**	10	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**
Total number of fastest detected cases		**9**	0	0	0	0	0	4	**10**	0	0	0	0	0	2	**10**	1	0	0	0	0	1
Legends : E=Euclidean C=City-block ND=no detection																						

From Table 5.1, the overall performances have shown that all the PCA models have effectively detected all of the specified abrupt fault cases investigated. Interestingly, all of the CMDS-PA methods (for all dimension and window size settings) are also successfully achieved the similar performance as equal to the linear PCA performance. Therefore, both methods can be viewed as performing equally in terms of detection capability with regard to abrupt fault cases. Comparing to the first CMDS method (Chapter 4), the results from CMDS-PA is relatively better, because not all of the sCMDS methods previously can produce 100% detection on those abrupt fault cases. Nonetheless, Table 5.2 also indicates that the best performance based on the total number of the fastest detection cases is still won by PCA, but the differences between PCA and CMDS-PA results are found hugely decreased as compared to sCMDS. In particular, 10 out of 11 detected cases are determined very efficient (through applying Euclidean distance and window size 5), which is 1 case lower than PCA. The generic trend of city-block results has also shown considerable improvement compared to sCMDS performance, but still not as good as that based on the Euclidean distance. All in all, the overall monitoring performance of CMDS-PA can be perceived as effective as well as efficient and equivalent to the overall PCA outcomes based on the abrupt fault cases.

Further improvements have also been observed from Table 5.2 which is for incipient fault performances. Almost all the CMDS-PA methods can perform 100% detection, excluding one method using the City-block distance. This can be regarded as one major achievement as the number of detected cases is greatly increased compared to the results of sCMDS based monitoring systems in the previous chapter. Besides, the total number of NDs is also considerably reduced compared to the first framework. In conjunction to this, the total number of fastest detection cases is also found slightly increased from those of sCMDS, but subjected to the monitoring system with Euclidean distance and window size 5. In analysing the same performance factor, CMDS-PA based monitoring system is also shown to be significantly superior to PCA based monitoring system. More importantly, PCA generally took more than 10 sampling times to detect the changes for many cases, whereas CMDS-PA (with Euclidean distance and window size 5) may merely apply less than 5 sampling times for detection, hence, a huge gap in terms of efficiency capability between the two can be clearly noticed. In analysing all of these trends, CMDS-PA can be viewed as to be performing better than PCA in terms of detecting the incipient faults.

The cases that specifically discussed in Chapters 3 and 4 are utilised again for detailed evaluation. Figures 5.3 and 5.4 show the monitoring statistic progressions for F6a and F6i respectively that specifically applied Euclidean distance and window size 5. Regarding F6a, Figure 5.3 shows that all the CMDS-PA models can efficiently detect the fault as early as after 1 sampling time (sampling time 3) from the time when the fault was introduced in the process (which is at sampling time 2) through both of the proposed monitoring statistics . This performance is also similar in comparison to those based on PCA and sCMDS that are discussed in Chapters 3 and 4 respectively. The monitoring outcomes of CMDS-PA for F6i are also found relatively analogous to sCMDS as indicated in Figure 5.4. In particular, the 3 dimensions model performed at equal rate of detection, while dimensions 5 and 7 have shown slightly quicker detection, which is one sampling time faster through S_r. Nonetheless, the performances of S_{m2} demonstrate poorer condition in relative to sCMDS, whereby slow detection periods are observed for all dimensions. Anyway, this particular drawback is compensated by the results of S_r, whereby both statistics work complementary. More importantly, all of these monitoring trends are significantly greater in performance when compared to the PCA results particularly based on F6i result.

The results of fault identification through utilising the contribution plot technique for F6a and F6i are presented in Figures 5.5 and 5.6 respectively. Regarding F6a (Figure 5.5), variables 7 (cooling water flow rate) and 13 (controller 2) have been found to be significant based on both types of contribution plot methods. In particular, the differential contribution plots have identified that particular trend initially by 2 sampling time earlier (sampling time 4) than those of conventional contribution plot results (sampling time 6) through S_{m2}. The trend has been found consistently enlarged until to the last sample by using both approaches. This fault leads to an increase in the cooling water flow rate (variable 7) which results in a reduction in reactor temperature. The reactor temperature control system then attempts to reduce the cooling water flow rate in order to maintain the reactor temperature by reducing controller 2 output (variable 13).

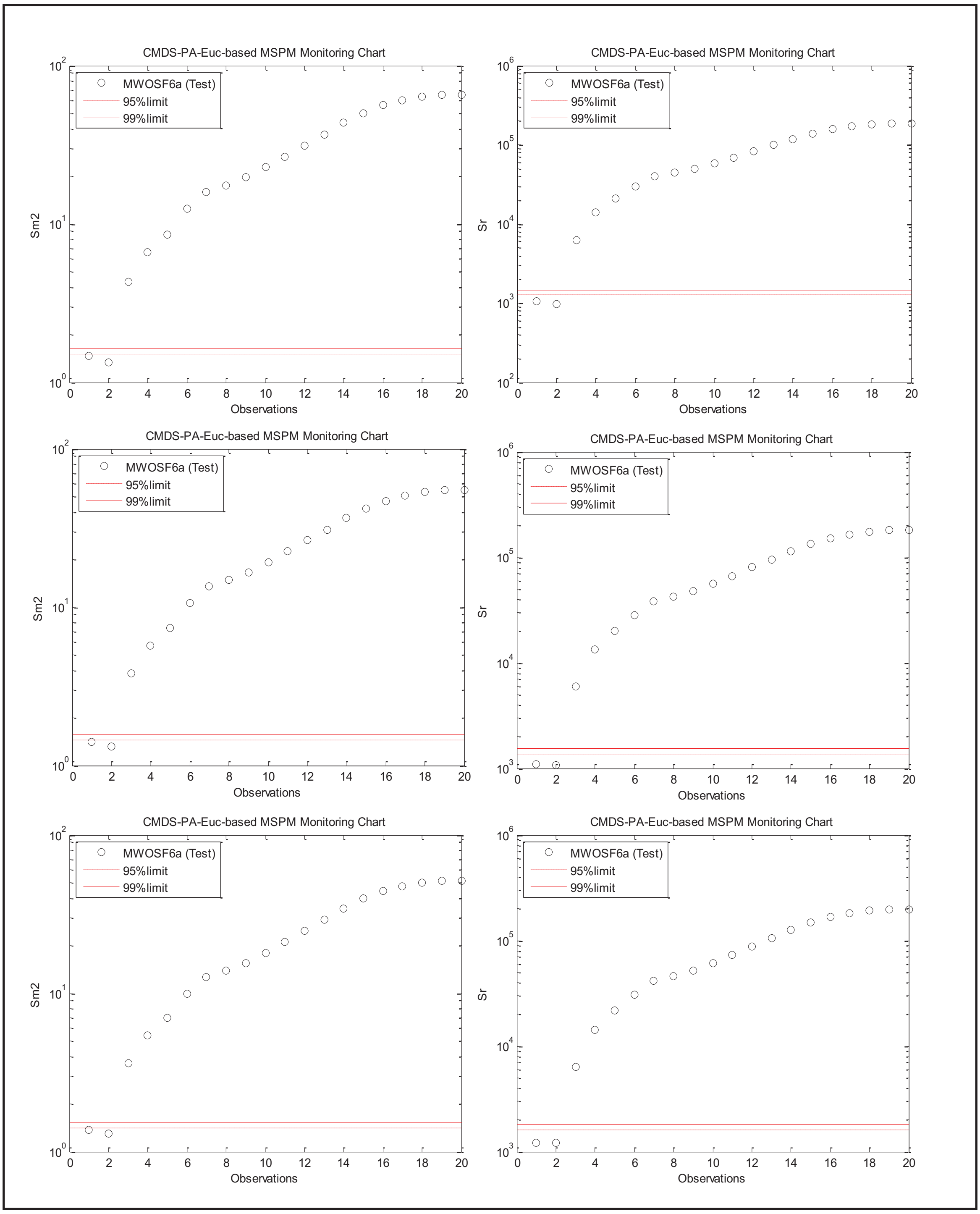

Figure 5.3: Monitoring progression of S_{m2}(left) and S_r (right) for F6a from monitoring systems based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

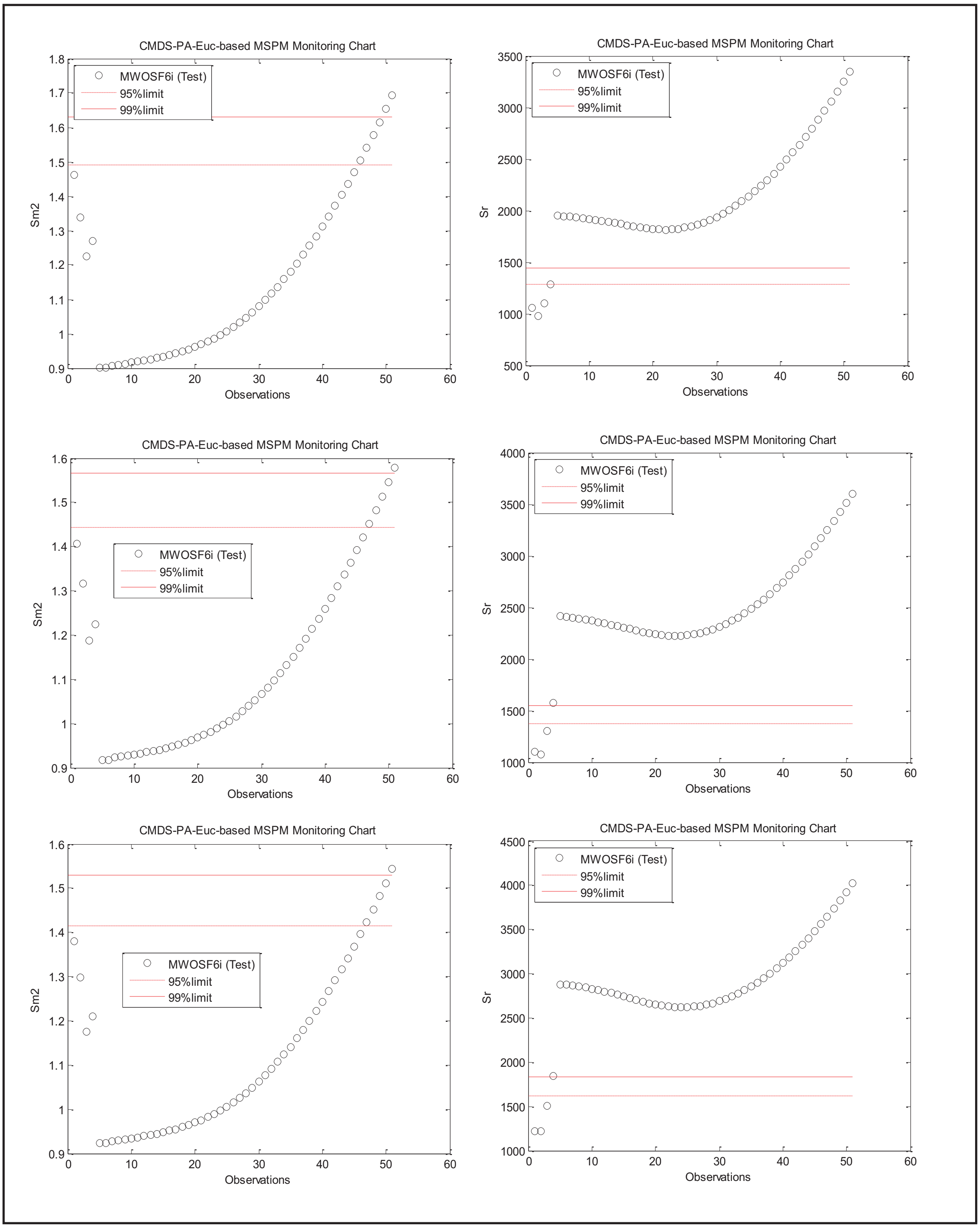

Figure 5.4: Monitoring progression of S_{m2} (left) and S_r (right) for F6i from monitoring systems based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

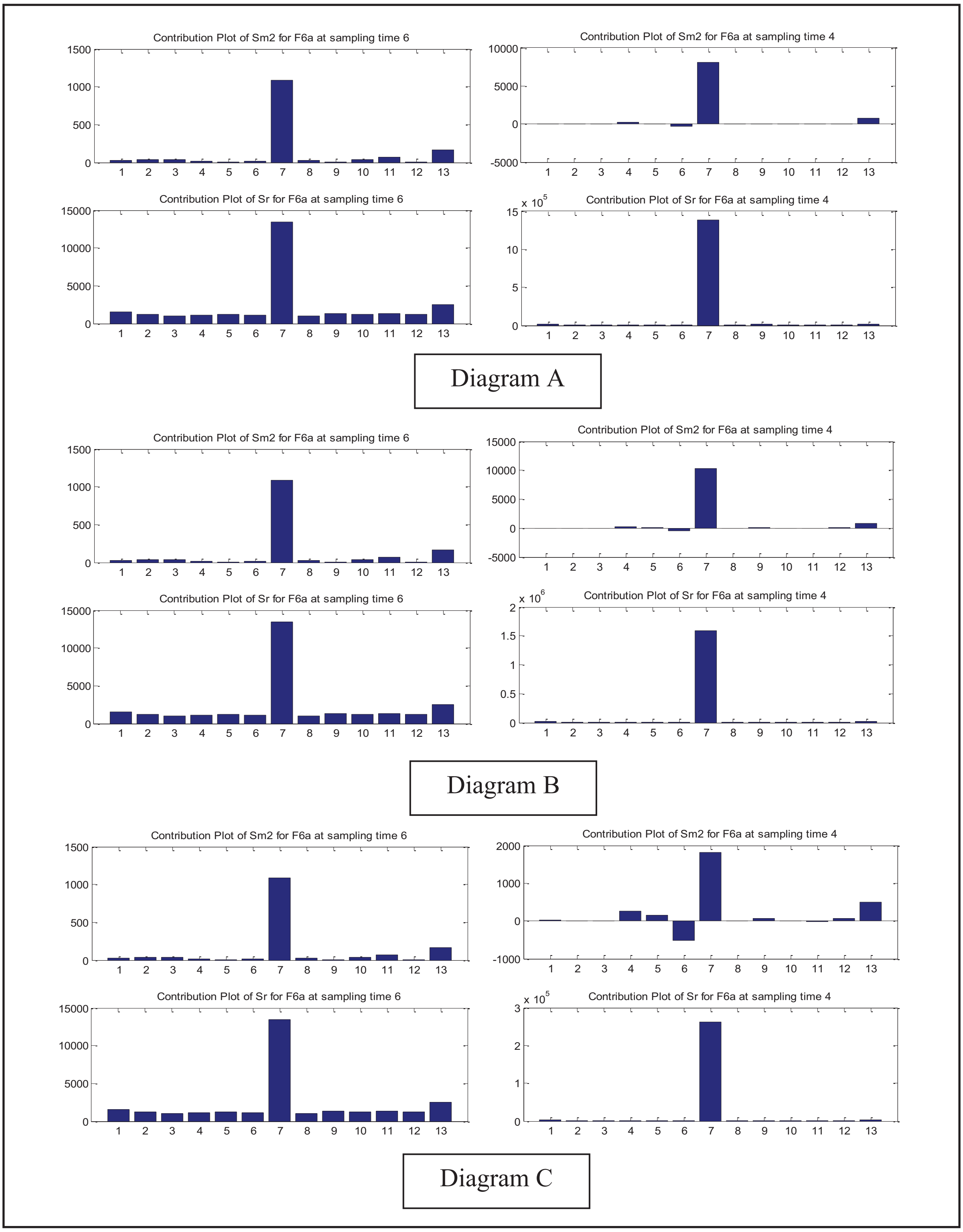

Figure 5.5: Conventional contribution plots (left) and differential contribution plots (right) for F6a from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C)

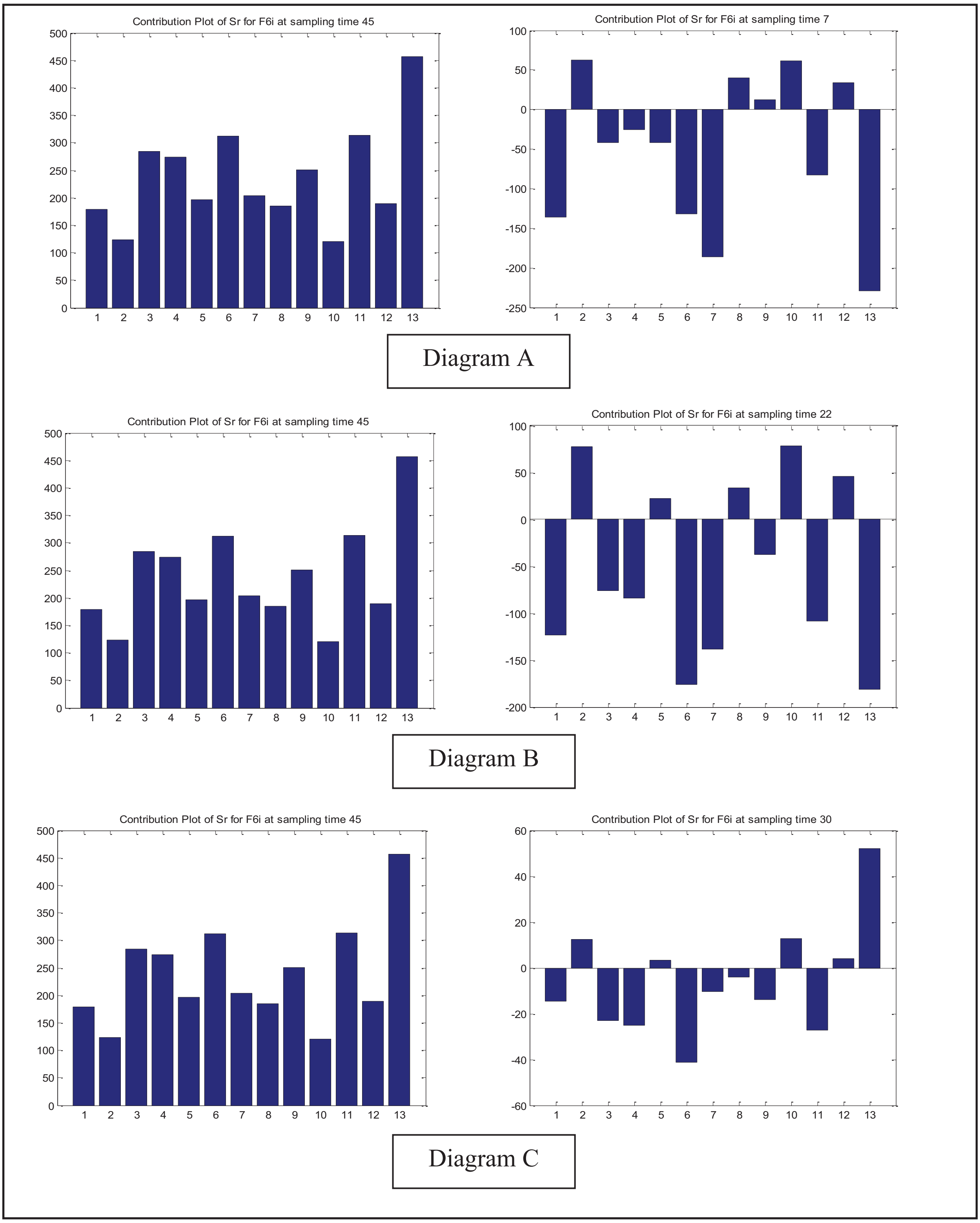

Figure 5.6: Conventional contribution plots (left) and differential contribution plots (right) for F6i from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C)

With respect to F6i (Figure 5.6), on the other hand, the results have shown that only variable 13 has been identified strongly contributing to the problem, and this was clearly indicated initially by the differential contribution plot at sampling time 7 (based on 3 dimensions and S_r statistic). Nonetheless, this particular magnitude of deviation has been observed increasing in downward direction instead of upward increment because the variable 13 score has been rearranged to be located far from the NOC variable location but still within the normal cluster initially. This trend has been observed consistent until sampling time 17 (the later samples have shown it started growing upwards until to the last sample). Other dimension applications have shown slower time of identification, but are still much quicker than the contribution plot outcomes (the first identification was seen at sampling time 45).

Figures 5.7 and 5.8 illustrate the trend of monitoring statistic progressions for F9a and F9i respectively of the CMDS-PA based monitoring system with Euclidean distance and window size 5. As similar to the performance of sCMDS, all models of CMDS-PA have detected F6a efficiently at sampling time 3 as indicated in Figure 5.7 through all monitoring statistics. The results are also comparable to the PCA performance with regard to the same fault data. Thus, CMDS-PA, sCMDS and PCA share similar outcomes based on F6a. Regarding F9i however, slight improvement over sCMDS has been noticed particularly relating to models with 5 and 7 dimensions as depicted from Figure 5.8. In particular, both of the models have detected F6i at 1 sampling time earlier (sampling time 4) than sCMDS (sampling time 5) based on S_r.

The results of fault identification using contribution plot technique for F9a and F9i are shown respectively in Figures 5.9 and 5.10. All the plots have indicated variable 12 (controller 3 – recycle flow control) that is undoubtedly linked to both F9a and F9i. In the case of F9a (Figure 5.9), both of the conventional and differential contribution plots have initially picked up the trend from sampling time 4 and this particular behaviour has been observed and sustained until the last sample. In particular, both statistics have signified the trend by using the conventional technique, whereas the trend was firstly observed by S_r through the differential contribution method (whereby both statistics have gradually indicated the deviation behaviour at the later samples). This abnormal condition is also believed to be related to variable 5 (which is also having a comparatively large in bar magnitude particularly based on F6a plots). All of these situations are occurred as a result of blockage that happened in the recycle stream pipelines. This fault will lead to a reduction in the recycle flow rate (variable 5). The recycle flow controller then attempts to increase recycle flow rate by increasing the controller output (variable 12). Compared to sCMDS, these identifications can be performed at earlier stage, which is 1 sampling time faster based on both statistics.

On the other hand, the clear identification for F9i (Figure 5.10) has to be performed at a much delayed period relative to sCMDS, specifically at sampling 40 based on S_r and using conventional contribution plot technique. However, the differential contribution plot results have produced better identification time than the conventional technique, whereby the bar corresponding to variable 12 has been noticed increased gradually starting from sampling time 25, and that is applicable for all dimension settings as well as through S_r.

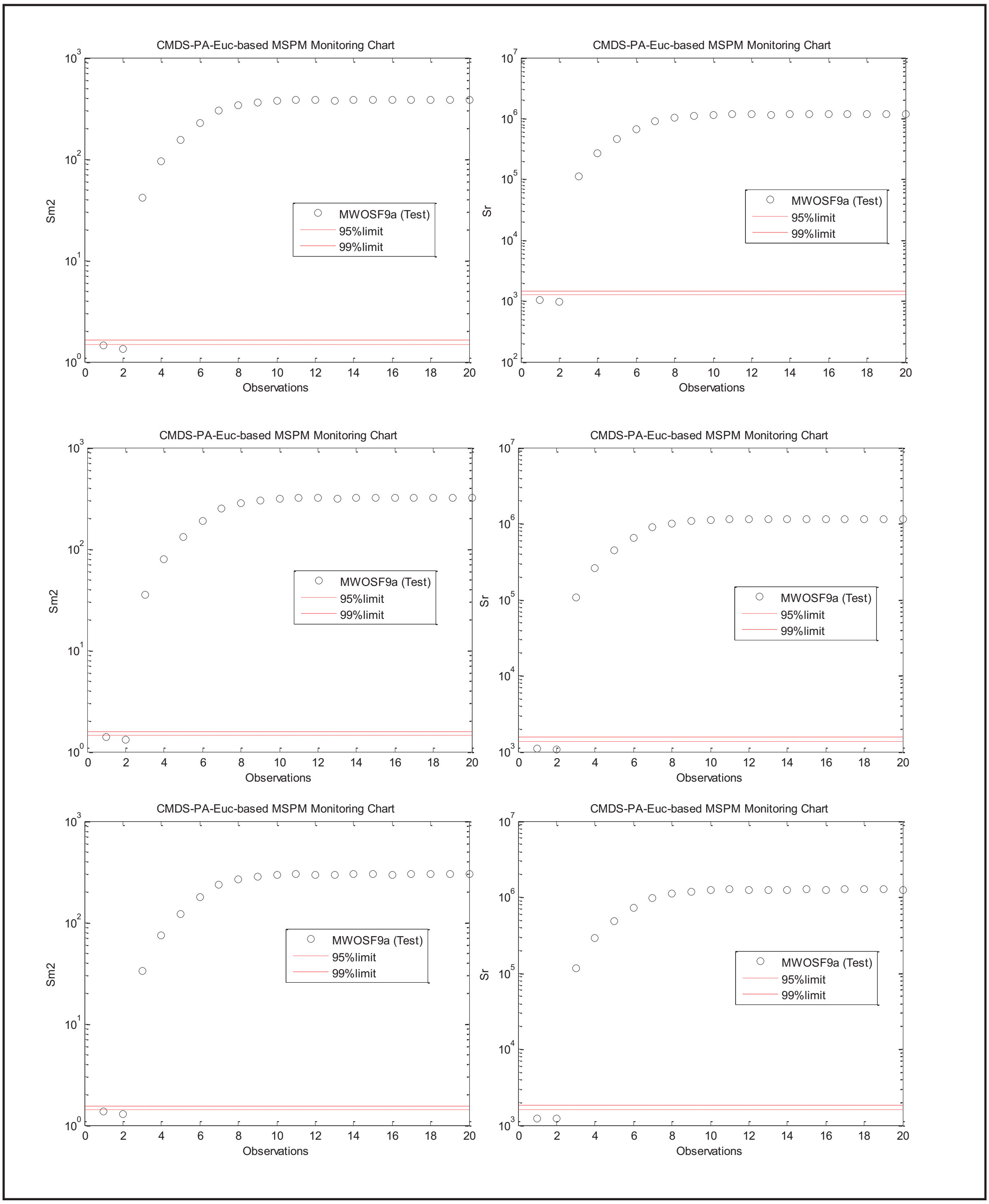

Figure 5.7: Monitoring progression of S_{m2}(left) and S_r (right) for F9a based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

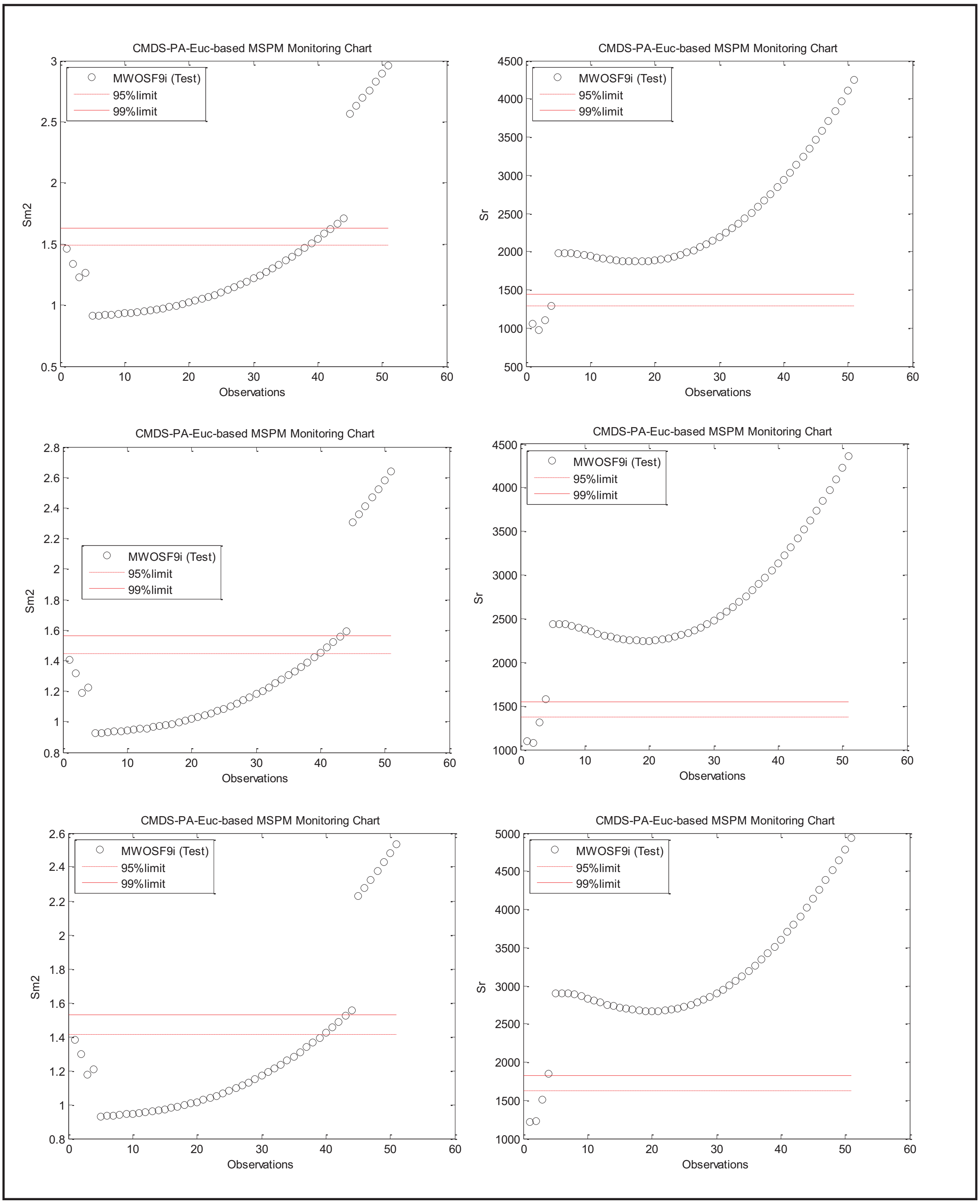

Figure 5.8: Monitoring progression of S_{m2}(left) and S_r (right) for F9i based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

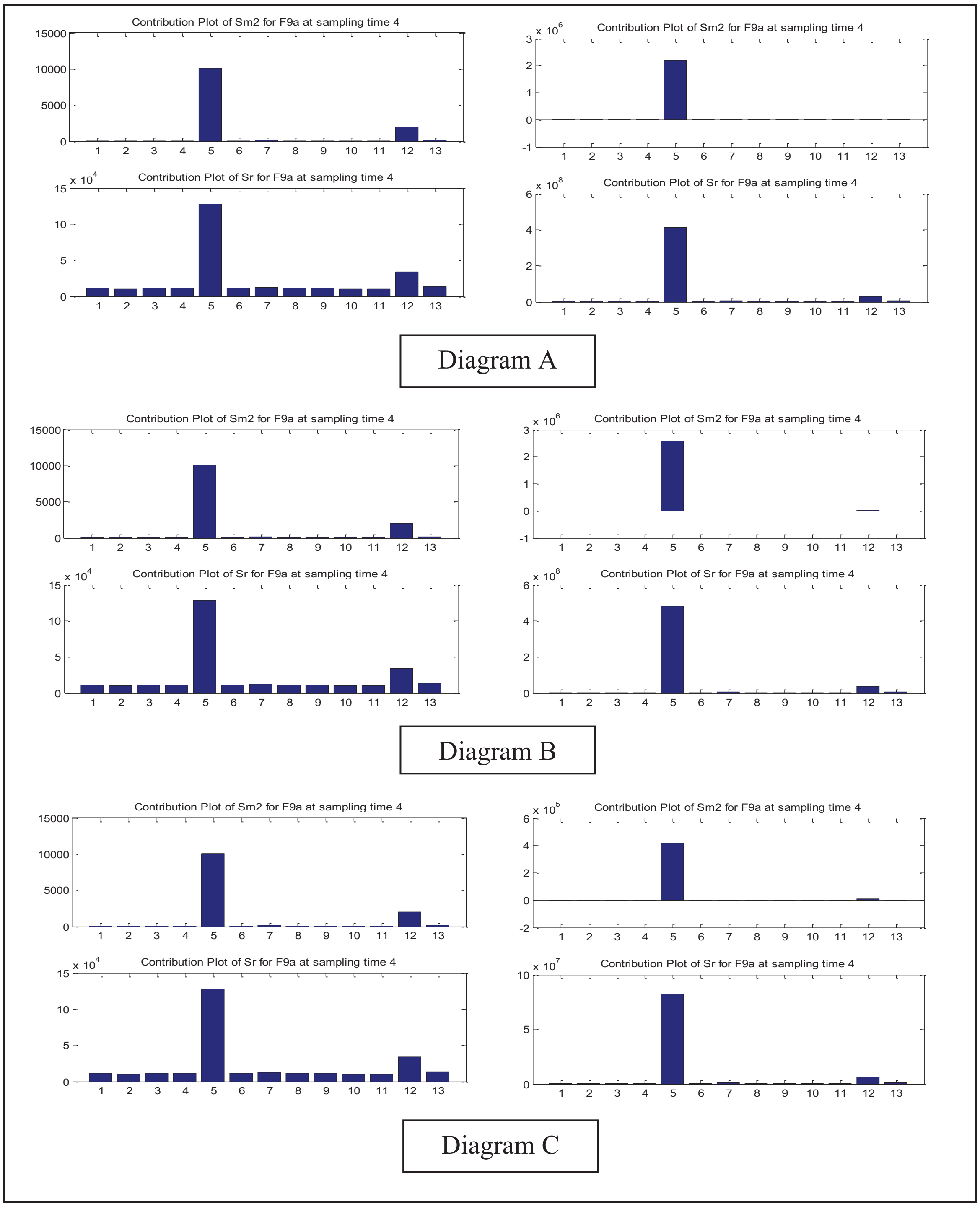

Figure 5.9: Conventional contribution plots (left) and differential contribution plots (right) for F9a from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C)

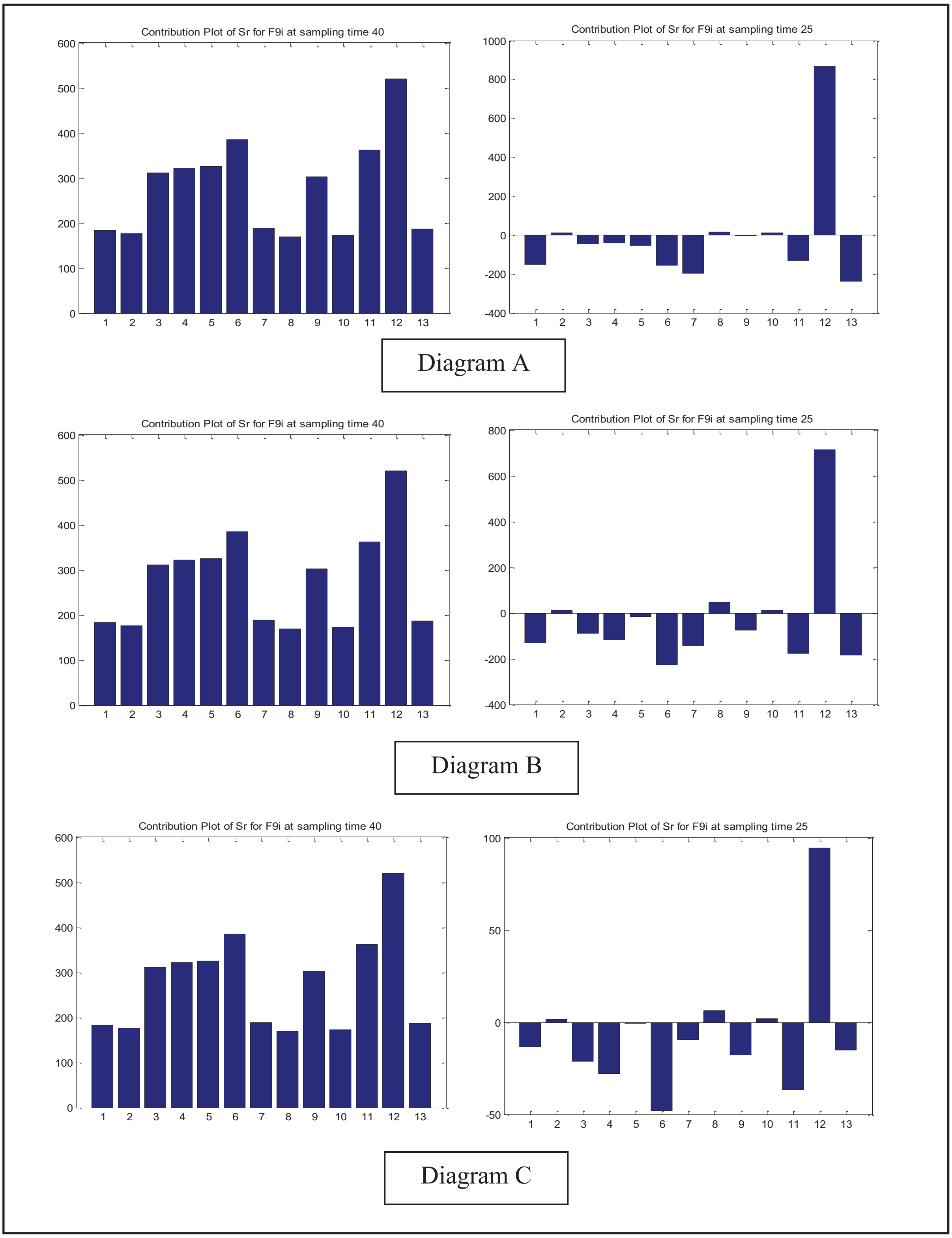

Figure 5.10: Conventional contribution plots (left) and differential contribution plots (right) for F9i from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C)

Finally, Figures 5.11 and 5.12 depict the monitoring outcomes for F11a and F11i respectively based on CMDS-PA based monitoring systems using Euclidean distance and window size 5. As can be observed in every other abrupt fault case previously, F11a can be also detected efficiently at sampling time 3 by all models as shown in Figure 5.11. The results are found absolutely analogous to those performances of sCMDS and PCA. The monitoring trends of F11i (Figure 5.12) are also determined to be almost identical in relation to sCMDS as well as PCA. In particular, this particular fault can be detected as early as at sampling time 5 (delayed by 3 sampling time from the time of fault introduction) through S_r.

Figures 5.13 and 5.14 summarised the performances of fault identification by utilisation of both contribution plot methods pertaining to F11a and F11i respectively. From both of the figures, all models have significantly specified variable 9 (inlet concentration) as the cause to both F11a and F11i through both conventional as well as differential contribution plot methods. In particular, both statistics have managed to identify F11a as early as at sampling time 3 (Figure 5.13). Meanwhile, the identification time of F11i was also found efficient, particularly at sampling times 7 and 8 using conventional and differential contribution plots respectively, but selective in terms of statistics (Figure 5.14). More specifically, the conventional and differential techniques have identified the abnormal behaviour through S_{m2} and S_r respectively for all dimensions. It is also noticed in further analyses that this particular abnormal condition has been progressively multiplied over the period of operation in every plot eventually. This finding strongly reflects to the true nature of responses of both cases that presented in Chapter 3, whereby, the concentration of inlet stream has been found either suddenly or gradually decreased and sustained over the time. In compared to sCMDS, CMDS-PA has been determined performing at equal rate for F11a , while in the case of F11i, the identifications of CMDS-PA have been found slightly slower.

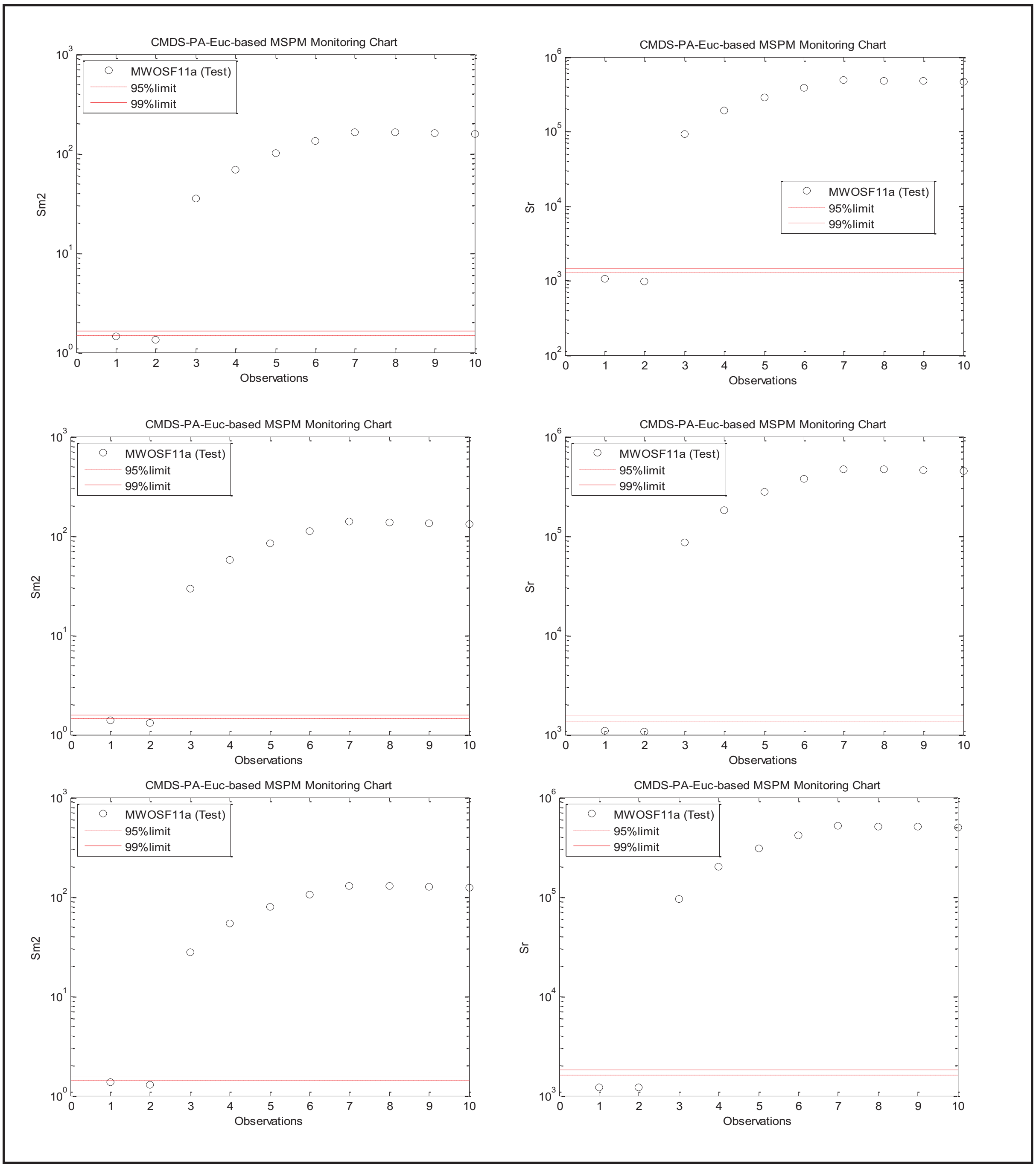

Figure 5.11: Monitoring progression of S_{m2}(left) and S_r (right) for F11a based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

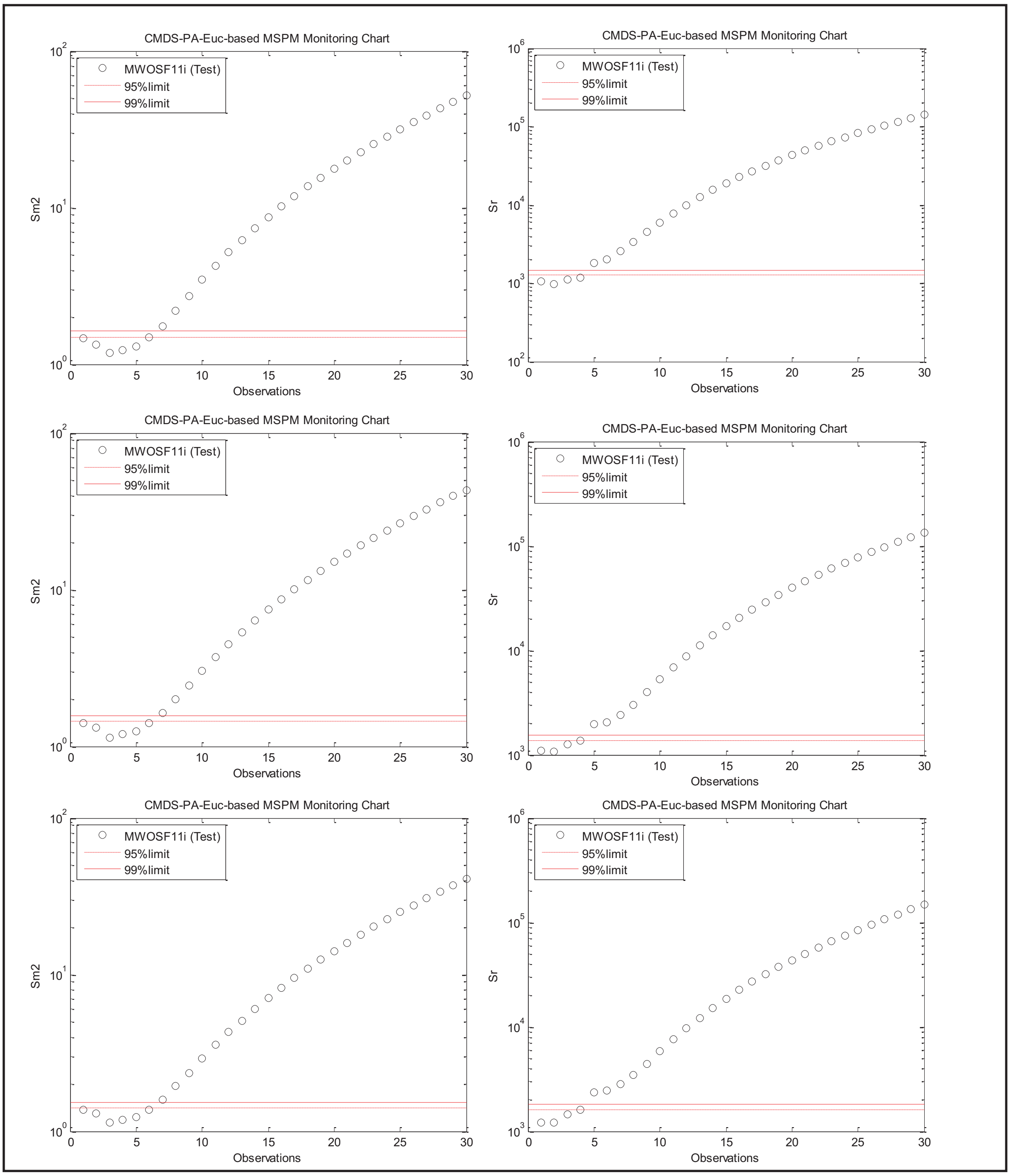

Figure 5.12: Monitoring progression of S_{m2}(left) and S_r (right) for F11i based on CMDS-PA using 3 dimensions (top), 5 dimensions (middle), and 7 dimensions (bottom)

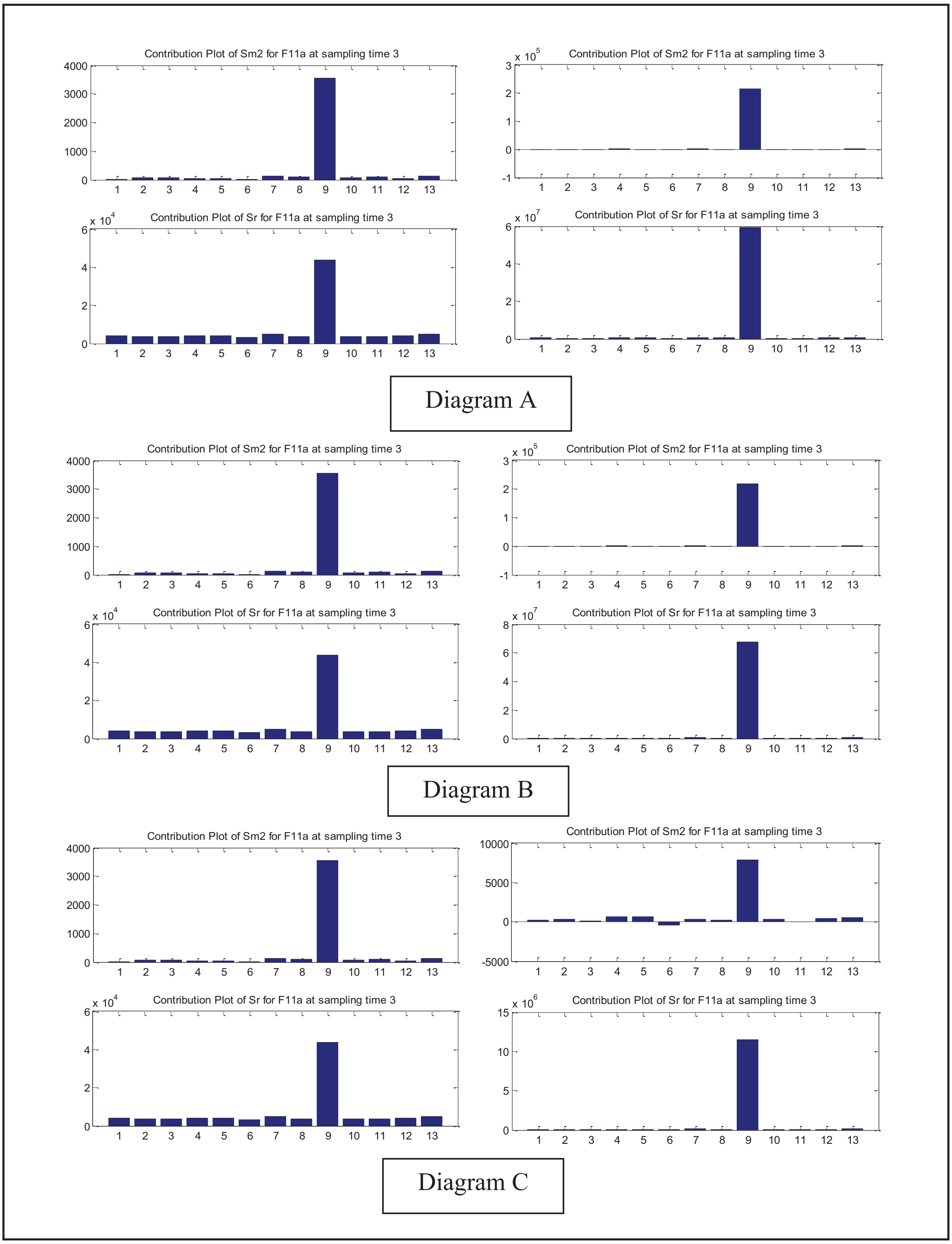

Figure 5.13: Conventional contribution plots (left) and differential contribution plots (right) for F11a from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C)

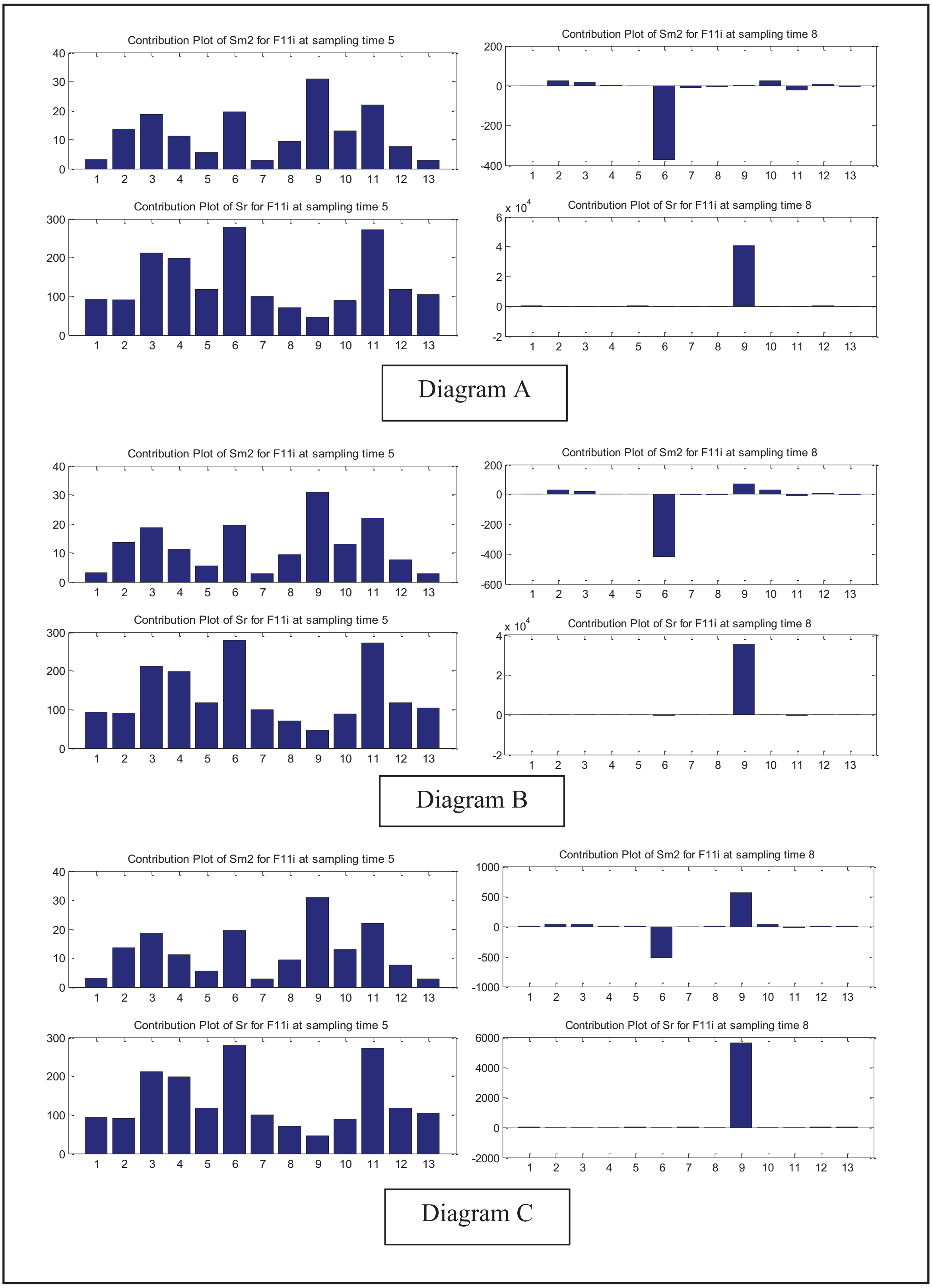

Figure 5.14: Conventional contribution plots (left) and differential contribution plots (right) for F11i from monitoring systems based on CMDS-PA using 3 dimensions (diagram A), 5 dimensions (diagram B), and 7 dimensions (diagram C)

5.4 Results Discussion

As similar to the previous discussion in Chapter 4, the objective of this section is mainly to explain the results of CMDS-PA from the perspective of the issues presented in the first chapter, where it also considers the perspective implication of the current method compared to the results of framework I.

5.4.1 The Impact of Dissimilarity Measures on The Monitoring Outcomes

This sub-section addresses the first question presented in Section 1.2 of Chapter 1. In Chapter 4, it was mentioned that Euclidean distance measure is advantageous compared to the City-block distance measure, and the results of CMDS-PA also holds the same suggestion, where this scenario can be obviously seen based on the number of fastest detection cases. However, this particular algorithm considers that the capability of monitoring system using City-block distance measure in detecting the fault is quite similar to that using the Euclidean distance measure regardless of the window sizes or dimensionalities applied.

5.4.2 The Impact of using New Monitoring Statistics on The Monitoring Outcomes

This sub-section is corresponding to the second question presented in Section 1.2 of Chapter I. As a result of integrating CMDS with PA, a different set of statistic pertaining to the magnitude of deviation has been used in contrary to sCMDS (as explained in the methodology section), while the concept of S_r is still valid. From the demonstration on the NOC cases of CSTRwR system, the overall performance was found very excellent, where all the FAR outputs were relatively low. Besides, all the specified malfunction events can be detected effectively as well as efficiently, which are comparable to PCA.

Table 5.3: Monitoring limits specified for S_r based on sCMDS and CMDS-PA methods

S_r Limits	sCMDS			CMDS-PA		
	Dimension 3	Dimension 5	Dimension 7	Dimension 3	Dimension 5	Dimension 7
95%	3930	4576	5038	1287	1374	1622
99%	4442	5137	5647	1439	1545	1829

However, this study has also found that the impact of using PA transformation functions can be clearly observed based on the S_r control limits. In particular, Table 5.3 shows the control limits for S_r calculated from both of framework I and II procedures previously, which are individually specified according to Euclidean distance and window size 5. The results denote that the control limits based on CMDS-PA for all models were significantly lower than those of sCMDS limits. Hence, this finding suggests that the inter-distance measures that projected by the standardized transformation factors of PA have effectively modified the sCMDS scores such that the new inter-distance measures (CMDS-PA scores) are much closer to the NOC1 dissimilarity scale configurations. In other words the squared errors in terms of dissimilarity scales between the scores of CMDS-PA and NOC1 has been reduced from those of sCMDS errors. This comparison is made exclusively based solely on S_r because the fundamental of S_{m1} is totally different from S_{m2} (although both statistics are used for representing the magnitude of deviation from a specified target).

5.4.3 The Impact of Applying Various Window Settings on The Monitoring Outcomes

This sub-section is trying to deliver the answer based on the third question presented in Section 1.2 of Chapter 1. It was explained in the previous chapter that by reducing the window size, it may contribute to initiate quicker in detection time. From the results of Table 5.2 on using Euclidean distance, this assumption has been obviously verified especially on the incipient fault cases, where great amount of abnormal situations have been detected much faster when using smaller window size. On the other hand, this assumption may not be that significant when using the City-block distance, where only those results of higher dimensions indicate such behaviour.

5.4.4The Impact of Applying Smaller Dimensionalities on The Monitoring Outcomes

This sub-section is pertaining to the last question presented in Section 1.2 of Chapter 1. In theory, increasing the number of dimensions will help basically to redevelop the CMDS scores such that they can emulate closely the original dissimilarity. Nonetheless, this does not imply that it will increase the capability of the system in detecting the fault as other factors are also significantly influencing the monitoring performance. In fact, increasing the dimensionality may also introduce more complexity on the system in maintaining the score redevelopment accuracy given that the number of variables is very large as illustrated in Table 5.3 previously. The results from Table 5.3 have denoted that the magnitudes of errors (limits) are almost linearly increased with the increment of dimensions. Thus, utilizing a reasonable lower dimensional-model for monitoring should be the primary objective (subject to the complexity of the system as well as the size of moving window selected).

In reflecting to the overall outcomes of CMDS-PA, even there were improvements in terms of performance factors that depicted in Tables 5.1 and 5.2 especially when more dimensions were used, the previous proposition on the dimensionality selection should be persistently held instead. Otherwise, the similar problem situation suffered by PCA will be also affecting the CMDS monitoring system applications. Thus, optimizing between model dimensionality, window size and number of variables should be further investigated, where a justified lower dimension model can be used for establishing an effective as well as efficient CMDS-based process monitoring system.

5.5 Summary

The main intention of this chapter is mainly to present the whole monitoring results based on the application of CMDS-PA approach. This method has been proposed in order to improve the sCMDS monitoring performance by standardizing the projection of the scores using PA transformation factors. In brief, the PA transformation factors have been formulated based on the correlation between two sets of CMDS NOC scores. Unlike sCMDS, these factors are used to modify the original scores in such a way the new scores can emulate closely the dissimilarity scales of NOC1. The overall results on the CSTRwR cases have shown significant improvements, where higher number of detected cases as well as faster detection has been achieved compared to both sCMDS and PCA. Furthermore, the number of NDs has been reduced greatly as compared to sCMDS. In general, both CMDS-PA and PCA have demonstrated equal performance based on abrupt fault cases, whereas CMDS-PA was found better in terms of sensitiveness (detection speed) than PCA in dealing with incipient fault cases.

CHAPTER 6

FRAMEWORK III: MDS-BASED MSPM SYSTEM USING MOVING WINDOW DYNAMICAL CMDS-PA PROJECTION

6.1 Introduction

This chapter presents the procedures as well as monitoring results based on the dynamical CMDS-PA-based MSPM system (CMDS-dPA), which is the last framework proposed in this study. In Chapter 4, the method focuses on developing the scores using exclusively the standard CMDS procedures. However, the method suffers from inconsistency in terms of score reproduction accuracy which leads to failure in detecting some of the specified abnormal cases (NDs) that are investigated. As a result, CMDS-PA was proposed (framework II), with the goal to enhance the mapping by applying PA procedures, and the outcomes have shown that those inconsistencies have been eliminated. Another important as well as positive implication was that the control limits for S_r have been reduced to some extent, which fundamentally means that closer mapping between the new and model configurations are obtained by way of dissimilarity scales.

However, the detailed analyses on some of the cases (such as F6i and F9i) have demonstrated slow detection particularly based on the S_{m2} monitoring statistic. Although this can be considered as slightly minor as this limitation can be directly complemented by using S_r (which are consistently efficient), but the gaps between the detection time of S_r and S_{m2} (referring to F6i and F9i) were found somewhat high. Thus, it has become the pure intention of this study to execute another algorithm in order to enhance the monitoring outcomes of CMDS-based system, specifically from this particular angle of perspective. In other words, the main aim of this particular framework is that, while it can maintain the FDT as to be considerable lower than those of linear PCA performances through S_r, but it must also attain faster detection on S_{m2}. It is believed that this can be performed through dynamic projection mechanism on the MWOS scores by the use of PA transformation procedures. This is simply achievable because the dynamic mapping may apply various sets of transformations factors, and this will provide great advantage to the CMDS method in modifying the MWOS scores to be configured as closely as possible to the NOC1 coordination. This will then, perhaps, reduce the errors between both sets of configurations, where much lower control limits can be established, and eventually, it can contribute perhaps to achieve quicker fault detection as well as identification.

This chapter is structured into four main sections, including the first as the introduction. The two subsequent sections present the detail methodology and the corresponding results respectively. The last section concludes this chapter.

6.2 Methodology

The fundamental approach of this framework is shown as in Figure 6.1.

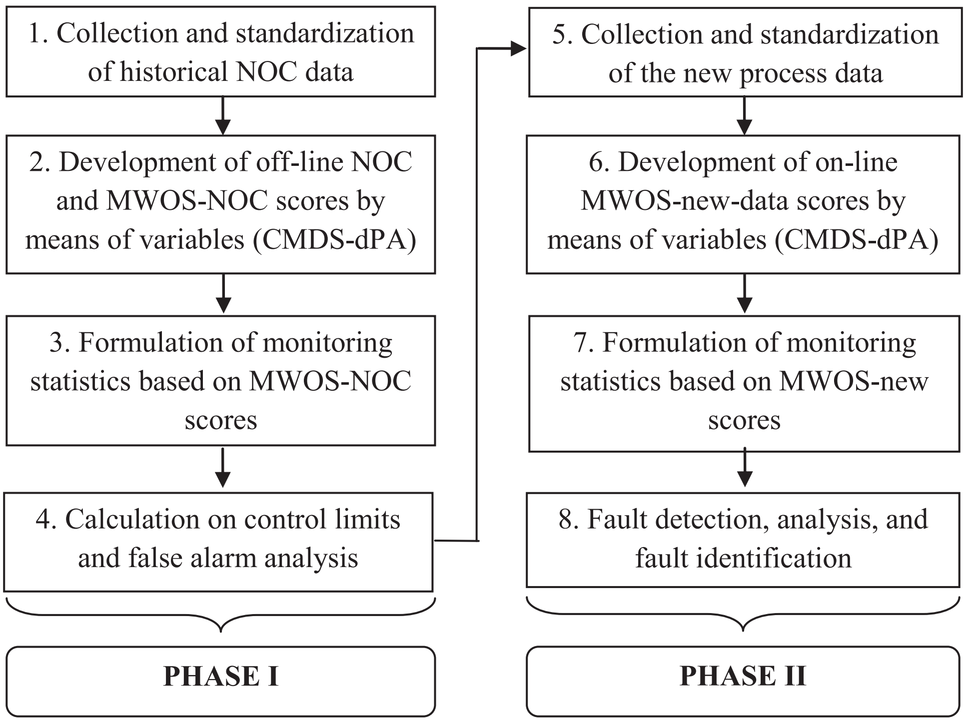

Figure 6.1. CMDS-dPA-based MSPM framework

Compared to framework I (Figure 4.1) and framework II (Figure 5.1), the main differences are on steps 2, 3, 6, and 7, where the projection of the scores will be executed by dynamic mapping through CMDS and PA applications at the same time. Unlike CMDS-PA, CMDS-dPA utilizing different sets of PA transformation factors in reconfiguring between the MWOS and NOC1 scores.

6.2.1 Phase I Procedures

In the beginning (**step 1**), the original NOC scores are divided into two sets, $\mathbf{X_{NOC1}}$ and $\mathbf{X_{NOC2}}$, after the standardization process. Then, in **step 2**, the standard CMDS algorithms are applied just on the first set of the NOC samples to represent the NOC score model, $\mathbf{Y_{NOC1}}$. Next, the moving window mechanism is applied integrating both of the NOC data sets in producing the MWOS-NOC samples, where various window sizes are adopted. Thus, there will be *'n'* number of samples of MWOS-NOC data produced at the end of the procedures. Each of these MWOS-NOC data will then has to apply the standard CMDS procedure in order to obtain the reproduction of the MWOS-NOC scores, $\mathbf{Y_{MWOS\text{-}NOC}}$. As the main intention is to obtain close approximation between two sets of score configurations, each of these MWOS-NOC scores will be used to get the unique sets of transformation functions individually by using the following PA procedures (Borg and Groenen, 1997):

i. Computation of the minor product moment between the first NOC scores, $\mathbf{Y_{NOC1}}$ (originated from $\mathbf{X_{NOC1}}$), and the MWOS-NOC scores, $\mathbf{Y_{MWOS\text{-}NOC(\mathit{k})}}$ (originated from $\mathbf{MWOS\text{-}X_{NOC(\mathit{k})}}$).

$$\mathbf{C_{PA(\mathit{k})}}=\mathbf{Y_{NOC1}}^T\mathbf{J_m}\,\mathbf{Y_{MWOS\text{-}NOC(\mathit{k})}} \quad (6.1)$$

where $\mathbf{J_m}$ is given in Equation 2.41 and $k = 1,2,....n$.

ii. Decomposition of $\mathbf{C_{PA}}$ into the eigen basic structures.

$$\mathbf{C_{PA(\mathit{k})}}=\mathbf{P_{PA(\mathit{k})}}\mathbf{V_{PA(\mathit{k})}}\mathbf{P_{PA(\mathit{k})}}^T \quad (6.2)$$

where $\mathbf{P_{PA(\mathit{k})}}$ and $\mathbf{V_{PA(\mathit{k})}}$ are the eigenvectors and eigenvalues matrices of $\mathbf{C_{PA(\mathit{k})}}$ respectively at MWOS sampling time k.

iii. Identification of the optimal rotation matrix:

$$\mathbf{R}_{(k)} = \mathbf{P_{PA(\mathit{k})}}\,\mathbf{P_{PA(\mathit{k})}}^T \quad (6.3)$$

iv. Identification of the optimal dilation scale:

$$s_{(k)}=(\mathrm{tr}\mathbf{Y_{NOC1}}^T\mathbf{J_m}\,\mathbf{Y_{MWOS\text{-}NOC(\mathit{k})}}\,\mathbf{R}_{(k)})/(\mathrm{tr}\,\mathbf{Y_{MWOS\text{-}NOC(\mathit{k})}}^T\mathbf{J_m}\,\mathbf{Y_{MWOS\text{-}NOC(\mathit{k})}}) \quad (6.4)$$

v. Identification of the optimal translation vector:

$$\mathbf{t}_{(k)} =(\mathbf{Y_{NOC1}}\text{-}s\,\mathbf{Y_{MWOS\text{-}NOC(\mathit{k})}}\,\mathbf{R}_{(k)})^T\mathbf{1}/m,\ m = \text{number of variables} \quad (6.5)$$

vi. The final equation for scores projection is given by:

$$\mathbf{Y_{dPA(\mathit{k})}}=s_{(k)}\,\mathbf{Y_{NOC\text{-}MWOS(\mathit{k})}}\mathbf{R}_{(k)}+\mathbf{1}\mathbf{t}_{(k)}^T \quad (6.6)$$

In contrary to framework II, each matrix of the MWOS-NOC samples in framework III produces different sets of PA transformation factors ($s_{(k)}$, $\mathbf{R}_{(k)}$, and $\mathbf{t}_{(k)}$) which are unique according to each of the MWOS-NOC matrix and this is where the term 'dynamic' is referring to. It is assumed that, the reconstructed $\mathbf{Y}_{\mathbf{dPA}(k)}$ will have a set of modified configurations which are significantly closer to $\mathbf{Y}_{\mathbf{NOC1}}$ relative to $\mathbf{Y}_{\mathbf{PA}(k)}$ from framework II (described in Chapter 5).

In **step 3**, framework III utilizes the same monitoring statistics as in framework II, which are S_{m2} and S_r. This is simply because both share the same monitoring assumption. Thus, those corresponding statistics for CMDS-dPA are provided in Equations 6.7 and 6.8 respectively.

$$S_{m2} = \sum_{i=1}^{m}\left(\sum_{j=1}^{p}\left[y_{dPA}(i,j) - y_{dPA-NOC1}(i,j) \right]^2 \right) \tag{6.7}$$

$$S_r = \sum\left(\sum_{i=1}^{m}\sum_{j=1}^{m}\left[d_{dPA(ij)} - \delta_{dPA-NOC1(ij)} \right]^2 \right) \tag{6.8}$$

In particular, both methods try to find a configuration that matches to the prescribed coordination of the original NOC scores (due to the application of PA) instead of depending on the dissimilarity measure (framework I). As similar to the first and second frameworks, these statistics are assumed to follow chi-squared distribution as in equation 4.3 (**step 4**), where false alarm rates (FAR) will be subsequently analysed.

6.2.2 Phase II Procedures

After standardization (**step 5**), the projection of the scores is also conducted dynamically according to the first phase procedure. In similar to **step 2**, moving window mechanism is applied in **step 6** to integrate both NOC and the new samples (MWOS-new samples), whereby different sizes of window sizes are applied. These MWOS-new samples will then individually execute the standard CMDS procedures and producing the MWOS-new scores, $\mathbf{Y}_{\mathbf{MWOS\text{-}new}}$. Then, each of these MWOS-new scores will find the PA transformation factors separately as described in the following procedures (Borg and Groenen, 1997):

i. Computation of the minor product moment between the first NOC scores, $\mathbf{Y_{NOC1}}$ (originated from $\mathbf{X_{NOC1}}$), and the MWOS-new scores, $\mathbf{Y_{MWOS\text{-}new(\mathit{k})}}$ (originated from $\mathbf{X_{MWOS\text{-}new(\mathit{k})}}$).

$$\mathbf{C_{PA(\mathit{k})}}=\mathbf{Y_{NOC1}}^T\mathbf{J_m}\,\mathbf{Y_{MWOS\text{-}new(\mathit{k})}} \quad (6.9)$$

where $\mathbf{J_m}$ is given in Equation 2.41 and $k = 1,2,....n$.

ii. Decomposition of $\mathbf{C_{PA}}$ into the eigen basic structures.

$$\mathbf{C_{PA(\mathit{k})}}=\mathbf{P_{PA(\mathit{k})}}\mathbf{V_{PA(\mathit{k})}}\mathbf{P_{PA(\mathit{k})}}^T \quad (6.10)$$

where $\mathbf{P_{PA(\mathit{k})}}$ and $\mathbf{V_{PA(\mathit{k})}}$ are the eigenvectors and eigenvalues matrices of $\mathbf{C_{PA(\mathit{k})}}$ respectively at MWOS sampling time interval-k.

iii. Identification of the optimal rotation matrix:

$$\mathbf{R}_{(k)} = \mathbf{P_{PA(\mathit{k})}}\,\mathbf{P_{PA(\mathit{k})}}^T \quad (6.11)$$

iv. Identification of the optimal dilation scale:

$$s_{(k)}=(\mathrm{tr}\mathbf{Y_{NOC1}}^T\mathbf{J_m}\,\mathbf{Y_{MWOS\text{-}new(\mathit{k})}}\,\mathbf{R}_{(k)})/(\mathrm{tr}\,\mathbf{Y_{MWOS\text{-}new(\mathit{k})}}^T\mathbf{J_m}\,\mathbf{Y_{MWOS\text{-}new(\mathit{k})}}) \quad (6.12)$$

v. Identification of the optimal translation vector:

$$\mathbf{t}_{(k)} =(\mathbf{Y_{NOC1}}\text{-}s\,\mathbf{Y_{MWOS\text{-}NOC(\mathit{k})}}\,\mathbf{R}_{(k)})^T\mathbf{1}/m,\ m = \text{number of variables} \quad (6.13)$$

vi. The final reproduction of scores for the new samples is given by:

$$\mathbf{Y_{dPA(\mathit{k})}}=s_{(k)}\,\mathbf{Y_{MWOS\text{-}new(\mathit{k})}}\mathbf{R}_{(k)}+\mathbf{1t}_{(k)}^T \quad (6.14)$$

Again, the fundamental difference at this stage between frameworks II and III is on the scheme in reproducing the scores, whereby previously it was conducted through one-off projection (as in the PCA projection score). In other words, the redevelopment of the scores of CMDS-dPA will apply different sets of transformation factors ($s_{(k)}$, $\mathbf{R}_{(k)}$, and $\mathbf{t}_{(k)}$) for different sets of MWOS samples as opposed to CMDS-PA (whereby only 1 set of transformations factors was used), Both of the statistics S_{m2} and S_r are calculated in **step 7** for each of the $\mathbf{Y_{dPA}}$ scores. This particular monitoring system is also assessed based on two main performance factors, (similar to the previous chapters), which are the total number of detected cases and total number of fastest detection cases.

As similar to framework II, this particular framework has also applied two types of contribution plots (conventional and differential) for fault identification. The first is shown as in Equation 6.15, while the second is presented by Equation 6.16.

$$\left(CMDS - dPA\ Statistics\right)_j = \sum_{i=1}^{m} \left(X_i\right)_j \tag{6.15}$$

Where, *(CMDS-dPA Statistics)*$_j$ = CMDS-dPA statistics (S_{m2} or S_r) at a particular sampling time '*j*'.

$(X_i)_j$ = contribution of the *i*th variable to CMDS-dPA statistics at a particular sampling time '*j*'.

$$\left(\mathbf{dc}\right)_j = \left[\left(Contribution_{CMDS-dPA(fault)}\right)_i - \left(Contribution_{CMDS-dPA(normal)}\right)_i\right]_j \tag{6.16}$$

In particular, **dc** (differential contribution) is a vector that contains '*m*' number of error magnitudes corresponding to those variables involved. Meanwhile, '*i*' is of the index for process variables, whereas '*j*' relates to sampling time. The term '$Contribution_{CMDS\text{-}dPA(fault)}$' refers to S_{m2} or S_r of those CMDS-dPA faulty samples, whereby the parameter of '$Contribution_{CMDS\text{-}dPA(normal)}$' associates to those of NOC statistics obtained through CMDS-dPA.

Both of these techniques of fault identification are executed once a fault is detected by the CMDS-dPA system. The analysis will be conducted from the first sample of detection until the trend of deviations is sustained over the period of operation. All of these steps were implemented in the CSTRwR case study that discussed previously in Chapter 3.

6.3 Results and Analysis

The same sets of NOC data used in the previous chapters were employed again in this approach. In particular, 3, 5 and 7 dimensional models were developed to represent around 75%, 90% and 99% proximity to the original dissimilarity scales. Besides, three window sizes, 5, 10 and 15 were selected. . The projection of the scores as well as monitoring statistics was executed according to the CMDS-dPA procedures explained earlier. The results of FAR on the original NOC as well as the testing NOC set have shown zero rates, which implied that the statistics were highly robust. Tables 6.1 and 6.2, present the overall FDT on the CMDS-dPA as well as PCA system based on models with 5 and 7 dimensions/PCs respectively. The condition on the fault detection scheme was also similar to the previous approaches, which is at least 3 successive samples of either monitoring statistics located over the corresponding 99% control limit.

Table 6.1: Results of FDT of abrupt fault cases based on CMDS-dPA for CSTRwR

FAULT DETECTION TIME (FDT)															
Fault Cases	**Dimensions**	**5**							**7**						
	Methods	**CMDS-dPA**						**PCA**	**CMDS-dPA**						**PCA**
	Window sizes	**5**		**10**		15			**5**		**10**		**15**		
	Scales	**E**	**C**	**E**	**C**	**E**	C		**E**	**C**	**E**	**C**	**E**	**C**	
1a	Fault Detection Time	**1**	ND	6	ND	13	ND	**1**	**1**	3	4	6	6	11	**1**
2a		2	ND	8	ND	ND	ND	**1**	2	3	5	8	10	16	**1**
3a		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
4a		**1**	2	2	3	3	4	**1**	**1**	2	**1**	3	2	4	**1**
5a		**1**	2	**1**	3	2	4	**1**	**1**	**1**	**1**	2	**1**	3	**1**
6a		**1**	**1**	**1**	2	**1**	2	**1**	**1**	**1**	**1**	2	**1**	2	**1**
7a		6	6	7	8	9	9	**2**	6	6	6	8	8	9	**1**
8a		2	2	4	4	5	5	**1**	2	2	3	5	4	7	**1**
9a		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
10a		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
11a		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Total number of detected cases		**11**	9	**11**	9	10	9	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**
Total number of fastest detected cases		8	5	6	4	5	4	**11**	8	6	7	4	6	4	**11**
Legends : E=Euclidean C=City-block ND=no detection															

Table 6.2: Results of FDT of incipient fault cases based on CMDS-dPA for CSTRwR

FAULT DETECTION TIME (FDT)															
Fault Cases	**Dimensions**	**5**							**7**						
	Methods	**CMDS-dPA**						**PCA**	**CMDS-dPA**						**PCA**
	Window sizes	**5**		**10**		**15**			**5**		**10**		**15**		
	Scales	**E**	**C**	**E**	**C**	**E**	**C**		**E**	**C**	**E**	**C**	**E**	**C**	
1i	Fault Detection Time	**3**	**3**	5	6	6	9	4	**3**	**3**	4	5	5	7	5
2i		**4**	5	8	9	11	12	7	**3**	4	6	7	8	10	8
3i		**2**	3	4	6	5	7	**2**	2	3	4	5	4	6	**1**
4i		14	**9**	16	12	18	14	25	11	**3**	6	8	12	16	15
5i		**8**	9	10	12	13	13	9	6	**3**	8	8	12	13	22
6i		8	11	**7**	14	13	14	34	**2**	3	6	8	11	13	21
7i		9	11	**7**	13	13	14	32	**2**	3	6	8	11	13	21
8i		**7**	10	10	16	18	19	24	4	**3**	6	8	11	13	15
9i		**4**	7	7	10	12	13	24	**2**	3	6	8	11	13	14
10i		**6**	7	7	10	12	13	25	**2**	3	6	8	11	13	16
11i		**3**	4	6	6	8	8	**3**	**3**	**3**	6	6	7	7	9
Total number of detected cases		**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**	**11**
Total number of fastest detected cases		**8**	2	2	0	0	0	1	**7**	5	0	0	0	0	1
Legends : E=Euclidean C=City-block															

As opposed to the FDT results presented in Chapters 4 and 5, Tables 6.1 and 6.2 only show monitoring results using 5 and 7 dimensions. The overall performances of using 3 dimensions were unsatisfactory. However, the results of FDT on abrupt fault cases that depicted in Table 6.1 have proved that both CMDS-dPA and PCA are equivalent in terms of detection capability, where both can achieve 100% detection. It is also observed that, PCA has maintained it consistent efficient performance by producing the fastest detection for all cases that detected, but nevertheless, the CMDS-dPA also has generated somewhat similar potential. Although, there were only 8 instances out of 11 detected cases that determined to be efficient, and that through applying settings of dimensions 5 and Euclidean distance, but the FDT delays (correspond to those cases that are considered as inefficient in relative to PCA – faults number 2a, 7a and 8a) between this particular CMDS-dPA scheme and PCA is too close, which is generally less than 5. In realizing this, both the new and conventional methods can be also perceived as equal in performance with regard to the second factor. Thus, the overall results of fault detection for abrupt fault cases have shown that CMDS-dPA and PCA are generally identical, and this finding is also in consistent to the nature of performances that demonstrated by the previous works (sCMDS and CMDS-PA).

The performances of CMDS-dPA on the incipient fault cases that presented in Table 6.2 also have signified interesting potentials. In particular all the CMDS-dPA methods have successfully detected all the malfunction events that analyzed in similar to the PCA achievement (100% detection). Nonetheless, CMDS-dPA has demonstrated superior performance over PCA with regard to the second performance factor, particularly by two means. Firstly, the total number of fastest detection cases of CMDS-dPA is considerably greater than the PCA results, and that through utilizing Euclidean distance and window size 5 (8 or 7 versus 1). Secondly, the FDT gaps between the two methods have been found very large (more or less than 20 sampling time) in many cases of the incipient fault category. Thus, the overall analysis based on the context of Table 6.3 has suggested that the performances of CMDS-dPA are relatively outstanding in compared to PCA. Unlike PCA, this generally indicates that CMDS-dPA is naturally advantageous especially for cases where the deviation of the faults grow slowly over the time of operation

In evaluating the results in rather great details, all the cases that are discussed individually in Chapters 4 and 5 are employed again. In particular, all of the results are based on applications that utilised Euclidean distance and window size 5. Firstly, Figures 6.2 and

6.3 depict the statistic progressions for F6a and F6i respectively. Figure 6.2 shows that F6a can be efficiently detected at sampling time 3, which is one sampling time after the fault occurred in the process. This performance is found in agreement to those performances of sCMDS as well as CMDS-PA. The results of F6i as indicated in Figure 6.3 also demonstrate an enhanced performance in relative to PCA. In particular, the best performance is shown by using CMDS model with 7 dimensions, where this particular fault can be detected as early as after 2 sampling time through S_r. More interestingly, the performance of S_{m2} also has detected F6i with an enhanced speed compared to the previous results of using CMDS-PA on a same basis.

The fault identification process has been performed through both conventional and differential contribution plots, but only the second was perfectly working for this particular framework. Figures 6.4 and 6.5 present the contribution plots for F6a and F6i respectively. The plots are corresponded to those applications that are shown in Figures 6.2 and 6.3 previously. In particular, all plots (both statistics) in Figure 6.4 have shown that variables 7 and 13 are importantly corresponded to F11a particularly started from sampling time 6 (the contribution plots based on earlier samples have indicated only on variable 7). This finding is almost comparable to CMDS-PA identification, even though it is delayed by two sampling time. Further evaluations on the later samples have indicated that the bar relating to variable 13 has increased gradually. This fault leads to an increase in the cooling water flow rate (variable 7) which results in a reduction in reactor temperature. The reactor temperature control system then attempts to reduce the cooling water flow rate in order to maintain the reactor temperature by reducing controller 2 output (variable 13).

The plots in Figure 6.3, on the other hand, have only indicated that variable 13 is significantly associated to F11i. This can be clearly observed starting from sampling time 40 and 42 respectively for models with 5 and 7 dimensions through both statistic progressions, whereby this particular trend was observed consistently increased over the time of operation in all plots afterwards. The results of earlier samples have shown no clear or permanent deviation pattern specifically. In comparing to the results of CMDS-PA previously, the CMDS-dPA outcomes have shown relatively slower identification. This is perhaps as a result of the impact of dynamic PA projection, whereby CMDS-dPA requires a great magnitude of deviations to clearly signify the specific deviation behaviour correctly.

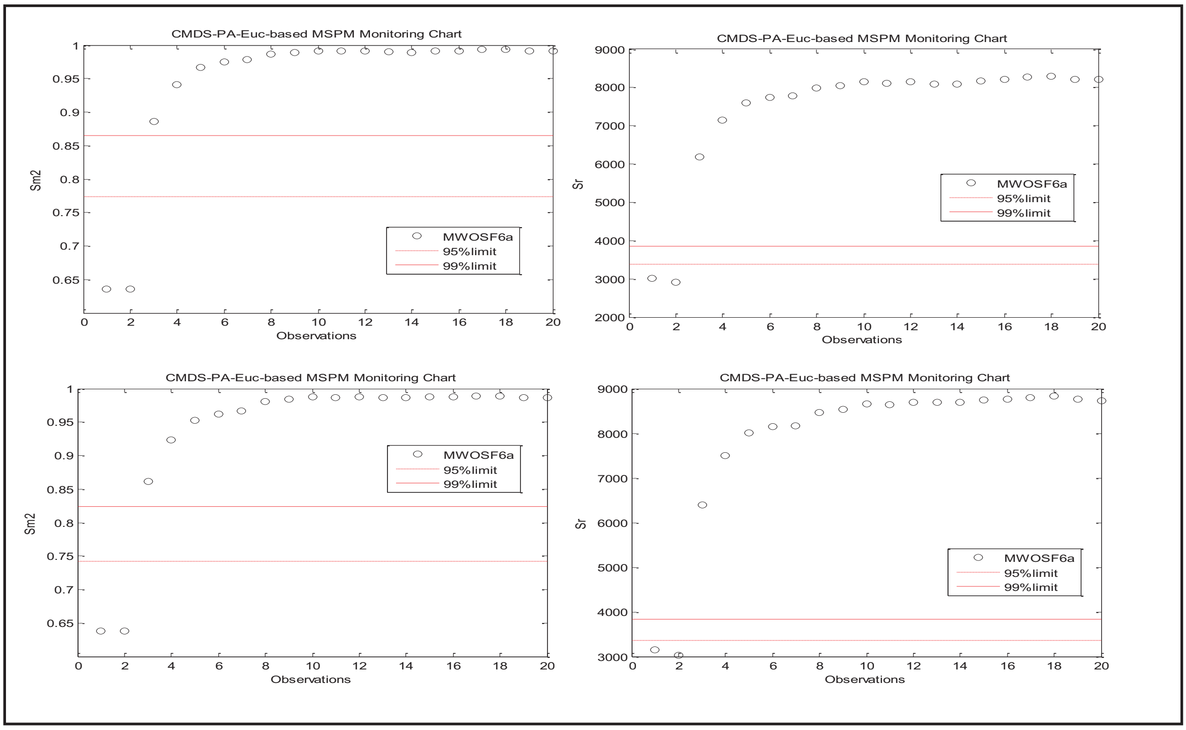

Figure 6.2: Monitoring progression of S_{m2}(left) and S_r (right) for F6a based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom)

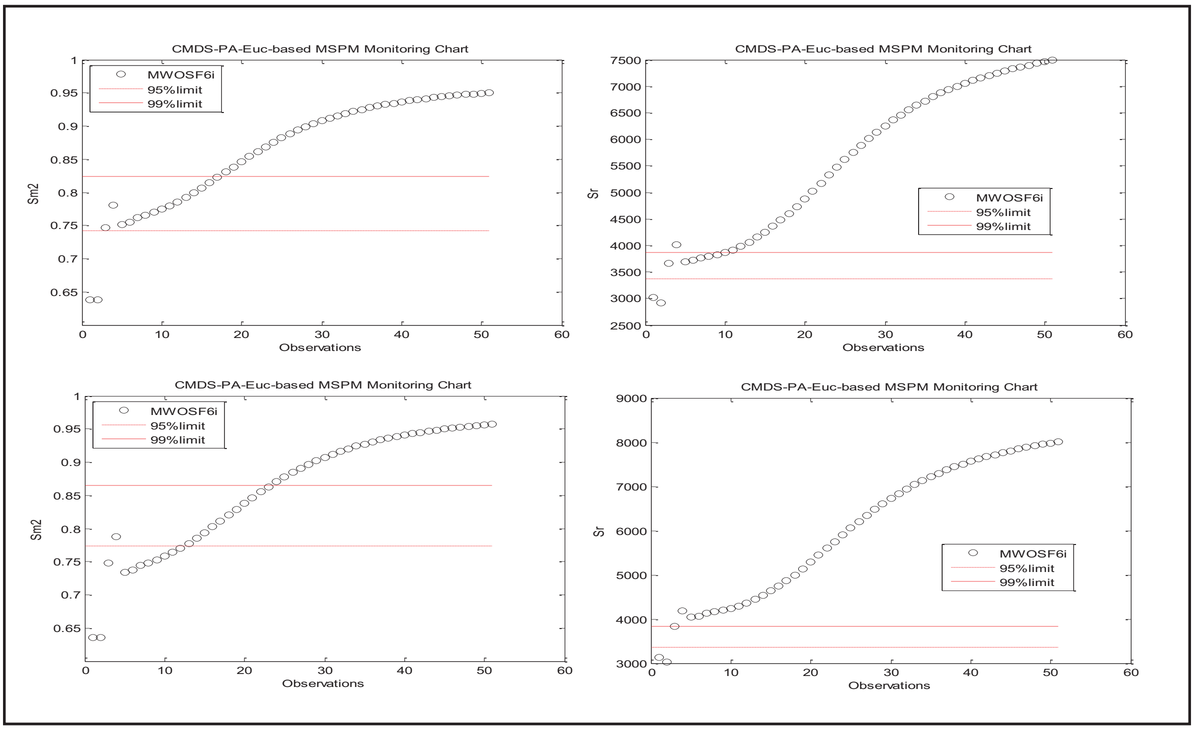

Figure 6.3: Monitoring progression of S_{m2} (left) and S_r (right) for F6i based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom)

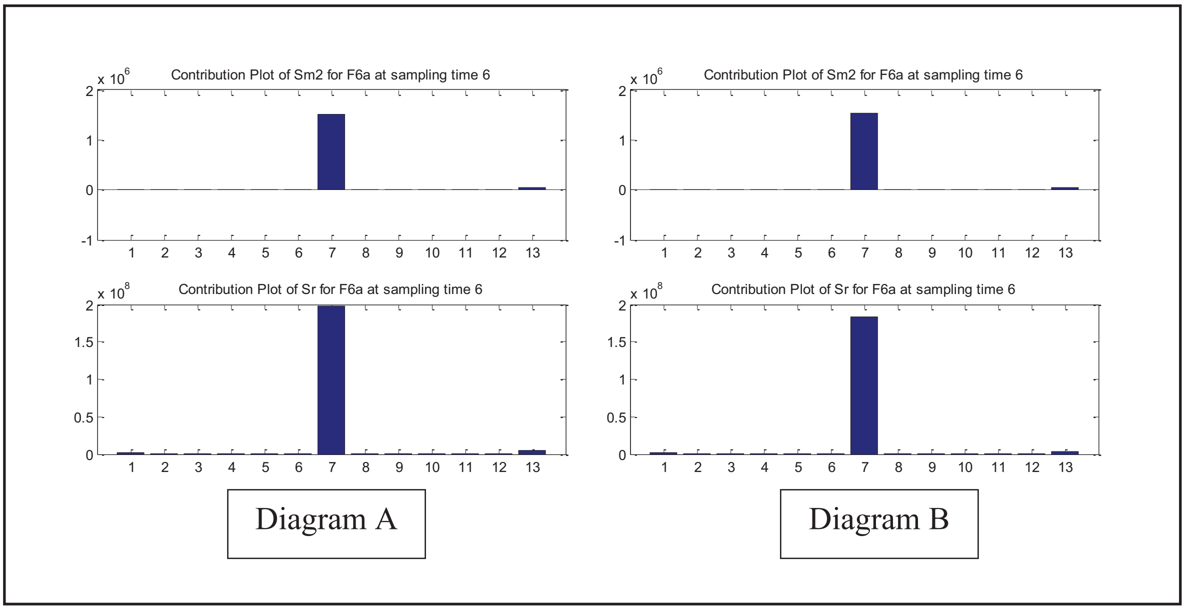

Figure 6.4: Differential contribution plots for F6a from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B)

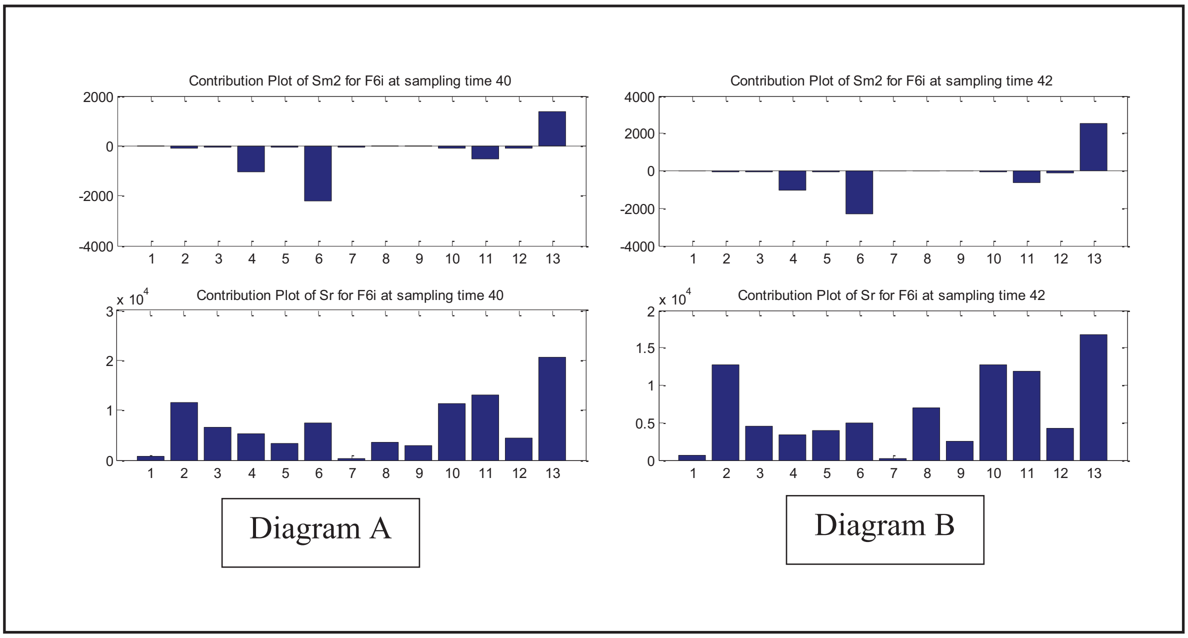

Figure 6.5: Differential contribution plots for F6i from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B)

The trends of statistic progressions based on F9a and F9i are shown in Figures 6.6 and 6.7 respectively which are from the CMDS-dPA approach using Euclidean distance and window size 5. In particular, the detection on F9a has been conducted efficiently as expected, which is at sampling time 3, through all the statistics that proposed as indicated in Figure 6.5. This result also has been found comparable to those monitoring performances from PCA, sCMDS and also CMDS-PA. The outcomes with regard to F9i also demonstrated intriguing performance, whereby the fault can be detected slightly faster by 1 sampling time in compared to sCMDS as well as CMDS-PA, and that through S_r and using dimension 7. It is also observed that the S_{m2} progressions for F9i have been also somewhat improved, whereby the detection can be executed in substantially shorter period in relative to CMDS-PA but longer than sCMDS. In contrary to PCA, this performance can be considered as excellent because the conventional method can merely detecting this particular fault generally more than 20 sampling time (refer to Table 6.3).

The results of fault identification were once again successfully conducted through differential contribution plot instead of conventional approach. The corresponding fault identifications of F9a and F9i are summarized in Figures 6.8 and 6.9 respectively. From Figure 6.8, all plots (both statistics) have indicated that variables 5 and 12 may have induced or affected strongly by F9a. The particular behaviour was observed started from sampling time 4 (sampling time 3 has clearly depicted solely on variable 5) and it has kept increasing (particularly on variable 12) over the time of operation. This fault will lead to a reduction in the recycle flow rate (variable 5). The recycle flow controller then attempts to increase recycle flow rate by increasing the controller output (variable 12).

Meanwhile, the contribution plots for F11i have only signified variable 12 as the sole contributor to the problem. This particular pattern was observed initially from sampling time 35 for both dimension applications as well as statistics (the results from earlier samples have shown no clear sign). The study has also learnt that the bar corresponding to variable 12 keeps enlarging linearly with the time of operation (corresponding to both statistics and dimensions applied). This observation is also found somewhat later than those of CMDS-PA identifications, and the reason could be related to the dynamic of PA projection that explained earlier.

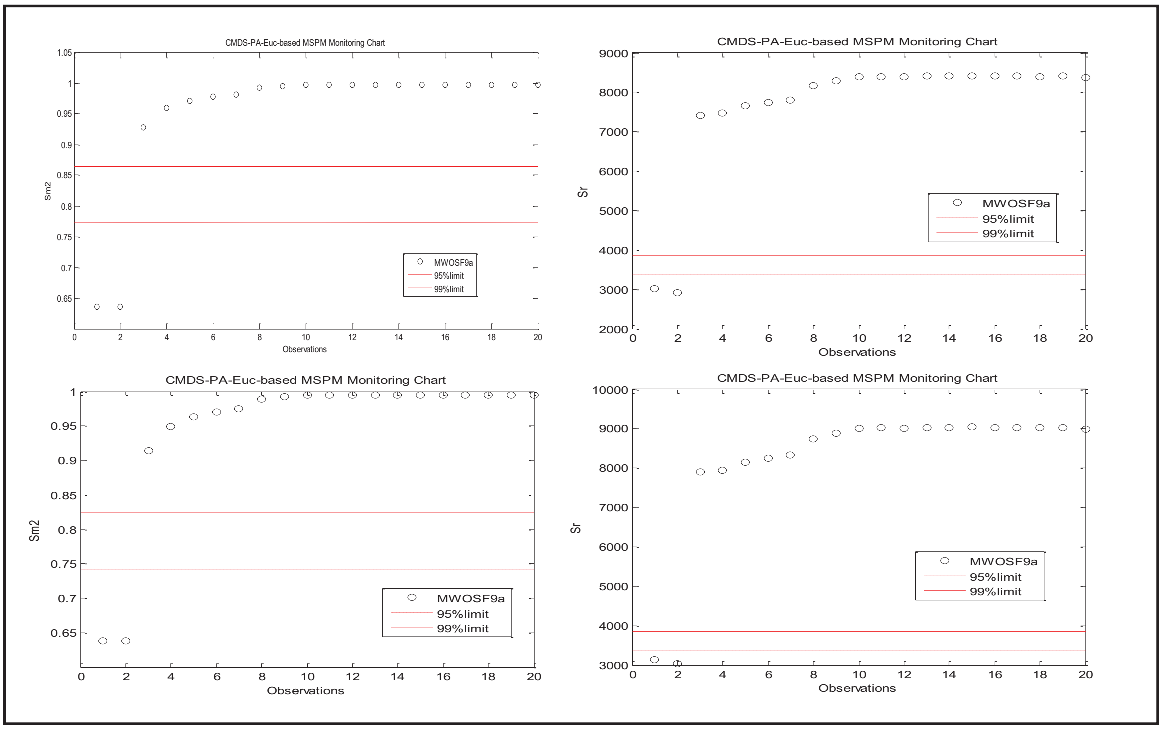

Figure 6.6: Monitoring progression of S_{m2}(left) and S_r (right) for F9a based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom)

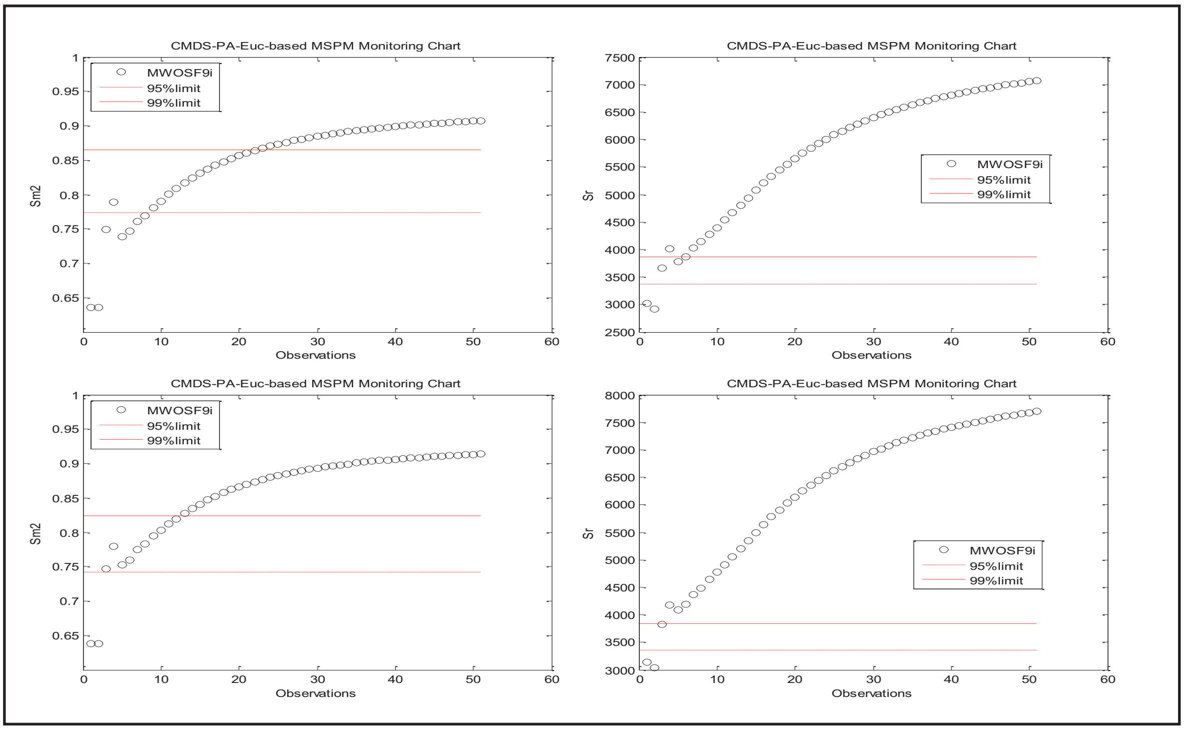

Figure 6.7: Monitoring progression of S_{m2} (left) and S_r (right) for F9i based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom)

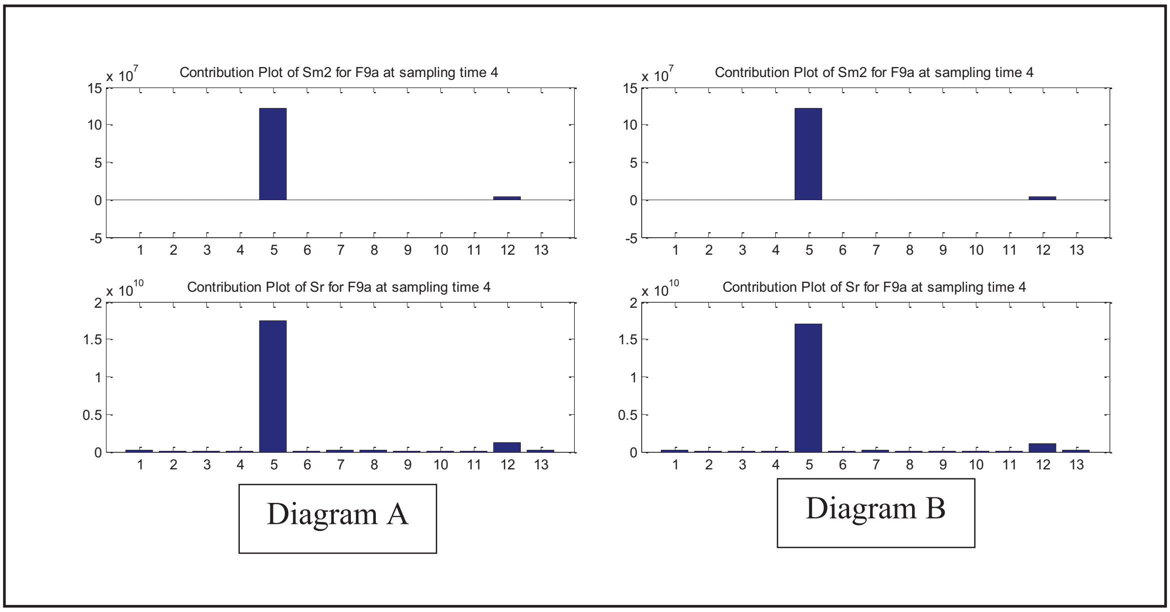

Figure 6.8: Differential contribution plots for F9a from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B)

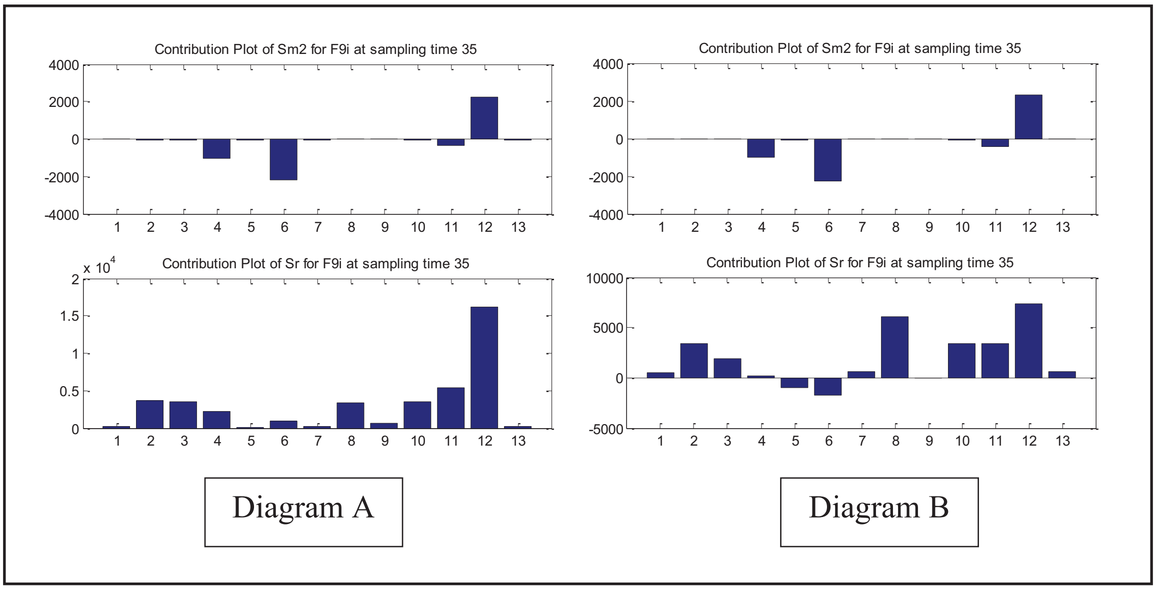

Figure 6.9: Differential contribution plots for F9i from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B)

Lastly, Figures 6.10 and 6.11 show the statistic progressions for F11a and F11i respectively. These results are based on using the Euclidean distance and window size 5. The results from Figure 6.10 have proved that CMDS-dPA is totally efficient, whereby F11a has been detected clearly at sampling time 3 through all the statistics that proposed. This performance is also identical to those performances of PCA, sCMDS as well as CMDS-dPA. In contra to the trend that denoted by previous abnormal cases, CMDS-dPA has demonstrated equivalent detection time, as illustrated in Figure 6.9, in relative to the previous CMDS monitoring systems with regard to F11i case. In general, the F11i event has been detected specifically at sampling time 5 through S_r. The progression of S_{m2} is also found almost comparable to the previous performances of CMDS methods.

The fault identification results were also successfully obtained through the differential contribution plots rather than conventional technique. The results are shown in Figures 6.12 and 6.13 for F11a and F11i respectively. In contra to the previous cases, the root of the F11a problem was identified very efficiently particularly at the instance of detection, which is at sampling time 3 through both statistics (and this particular behaviour was also seen sustained throughout the operation in the later samples). More specifically, all plots have indicated variable 9 (inlet concentration) as having the largest bar, which is obviously connected to F11a. Meanwhile, the identification for F11i was also somewhat efficient, which is at sampling time 8 (3 sampling time delayed from the time of detection), also through both statistics. In particular, the bar connecting to variable 9 starts to be the largest among the others, and it stays on growing during the later samples. The result is also generally comparable to those of CMDS-PA performances based on this particular case.

From the analysis of fault identification on the three specified cases, it is realized that framework III has a unique characteristic against the second CMDS framework. In general, the conventional contribution plot has failed to recognize the difference between the abnormal operation samples with the original NOC model based on the CMDS-dPA method. This is perhaps because the dynamic projection performed by PA is strongly modifying the outlier location (fault variables), such that, the new configurations are transformed and matched as exactly as possible to the original NOC coordinates. Fortunately, the outcomes of differential contribution plot technique have successfully identified the changes, but with a rather slower period compared to CMDS-PA.

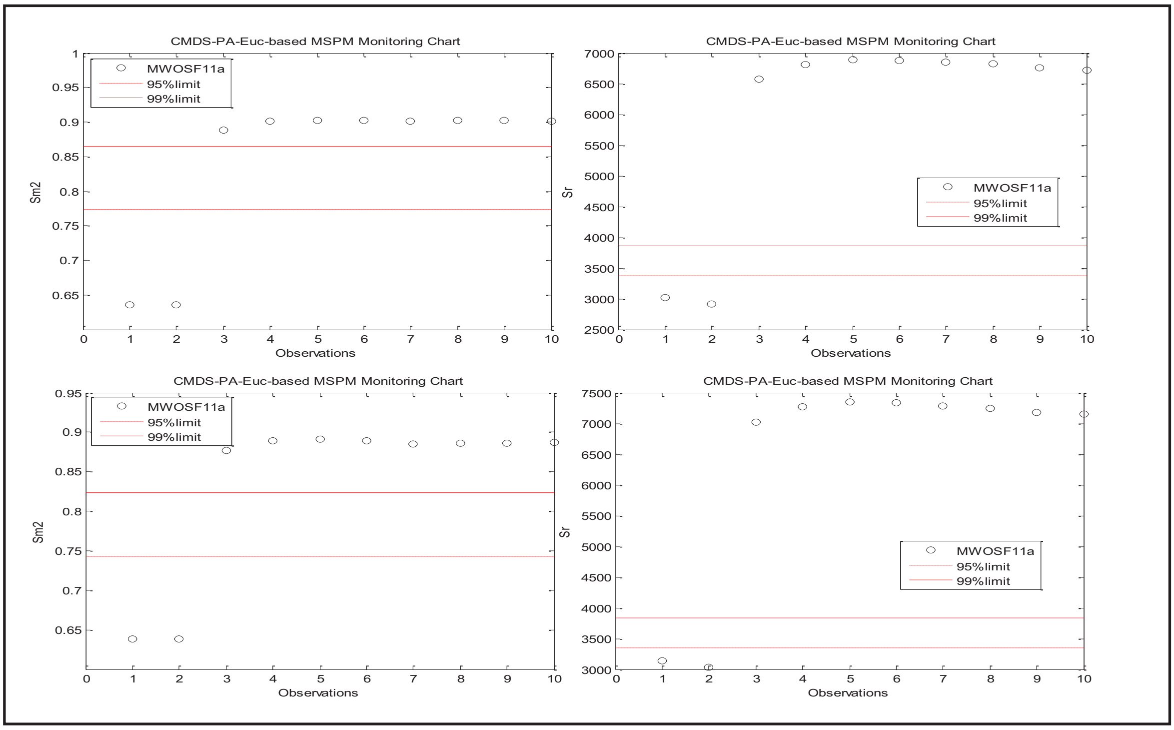

Figure 6.10: Monitoring progression of S_{m2}(left) and S_r (right) for F11a based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom)

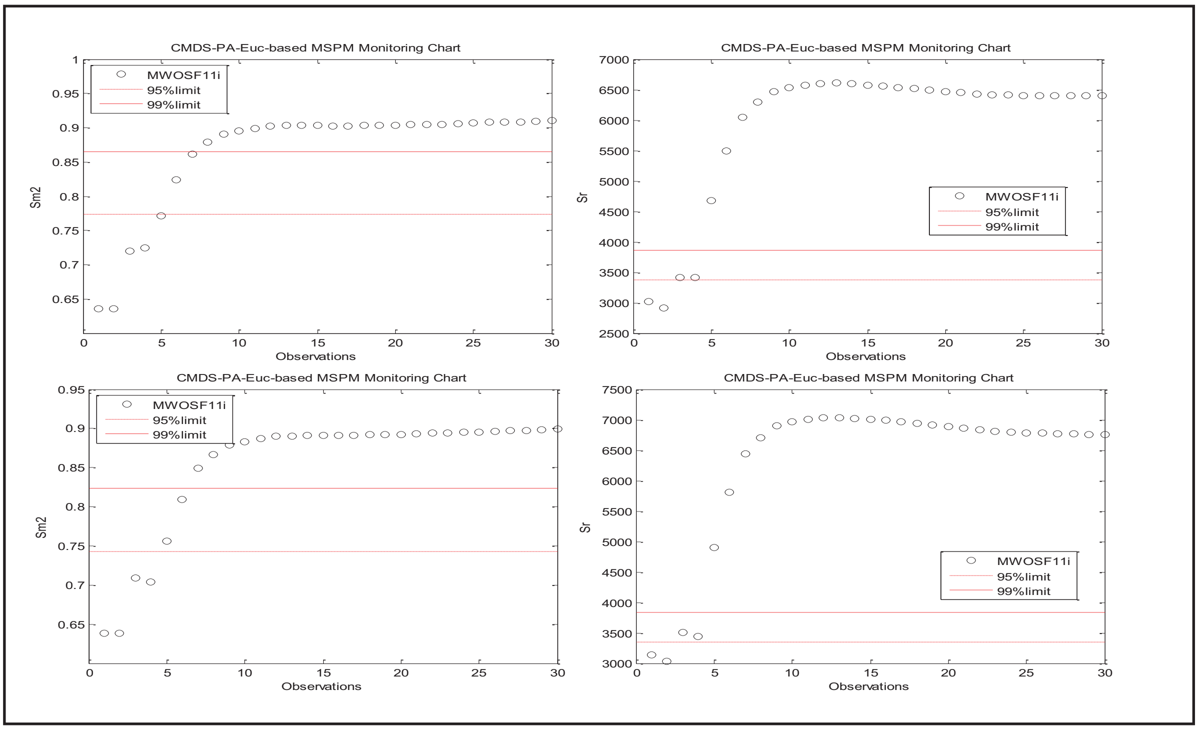

Figure 6.11: Monitoring progression of S_{m2}(left) and S_r (right) for F11i based on CMDS-dPA using 5 dimensions (top) and 7 dimensions (bottom)

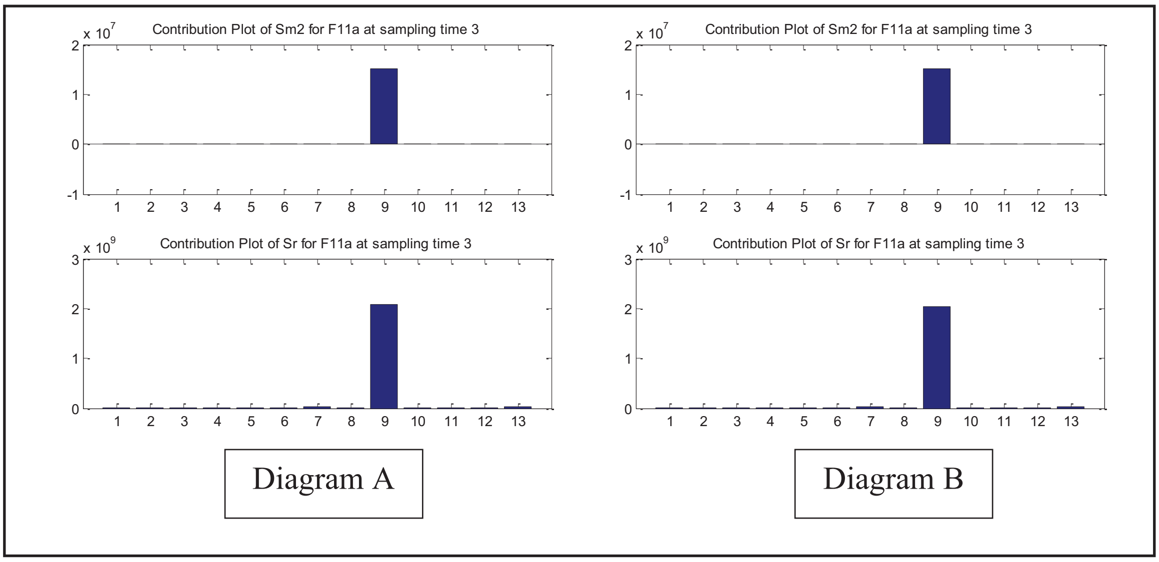

Figure 6.12: Differential contribution plots for F11a from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B)

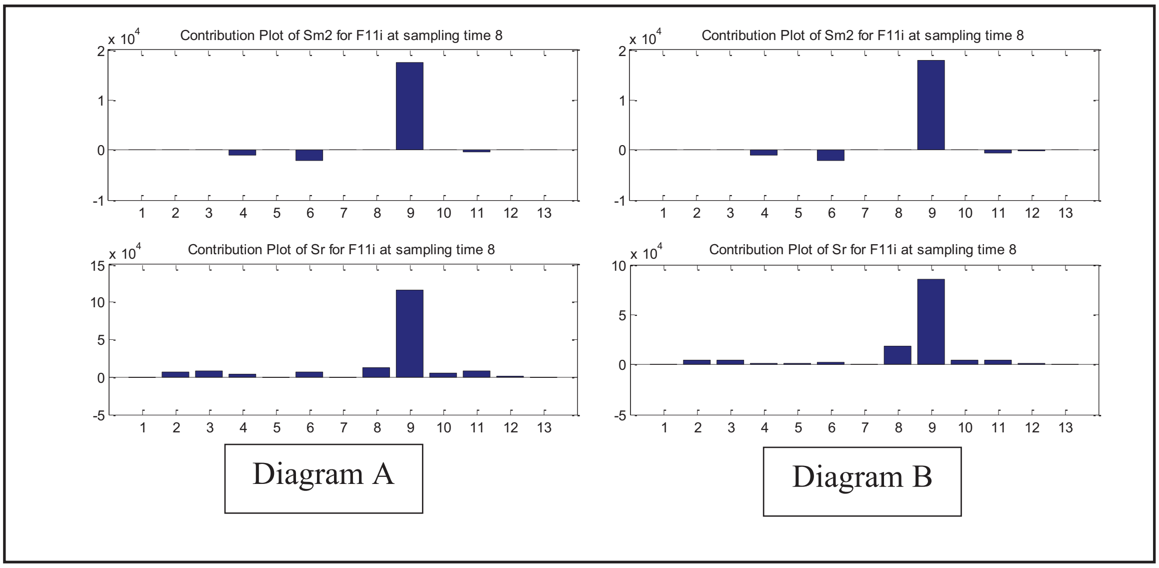

Figure 6.13: Differential contribution plots for F11i from monitoring systems based on CMDS-dPA using 5 dimensions (diagram A) and 7 dimensions (diagram B)

6.4 Results Discussion

6.4.1 The Impact of Dissimilarity Measures on The Monitoring Outcomes

This sub-section addressing the first question presented in Section 1.2 of Chapter I. From the previous results, Euclidean distance measure has been shown to be advantageous over City-block distance measure. The overall results based on the CMDS-dPA also indicate the same finding especially based on the FDT rates results. Thus, the original perspective is still preserved.

6.4.2 The Impact of using New Monitoring Statistics on The Monitoring Outcomes

This sub-section is corresponding to the second question presented in Section 1.2 of Chapter I. One major drawback identified from framework I was that there were a number of occurrences with regard to ND events. Thus, one of the speculated factors which believed contributed to the problem was based on the inconsistency of the sCMDS algorithms in projecting the scores (because it uses different eigenvector settings in the projection procedures). This effect was then removed or smoothened by the use of PA transformation factors in framework II. In particular, all of the scores (regardless of their individual eigenvector settings) are enforced to be standardized through the PA procedures and the results indicated that the squared errors in terms of dissimilarity scales were greatly reduced. Although the overall results have demonstrated major improvement, the detail analysis has indicated minor limitation, and that is slow progression especially on S_{m2} as well as fault identification. Thus, this particular framework was proposed to enhance the capability of CMDS on that particular aspect.

The control limits calculated for S_{m2} and S_r of CMDS-PA and CMDS-dPA procedures are shown in Table 6.3 and 6.4 respectively.

Table 6.3: Monitoring limits specified for S_{m2} based on Euclidean scale and window size 5

Sm2 Limits	CMDS-PA		CMDS-dPA	
	Dimension 5	Dimension 7	Dimension 5	Dimension 7
95%	1.44	1.42	0.77	0.74
99%	1.57	1.53	0.86	0.82

Table 6.4: Monitoring limits specified for S_r based on Euclidean scale and window size 5

Sr Limits	CMDS-PA		CMDS-dPA	
	Dimension 5	Dimension 7	Dimension 5	Dimension 7
95%	1375	1622	3373	3355
99%	1545	1829	3861	3841

It is obviously shown from Table 6.3 that the control limits or the squared errors in terms of resultant vector lengths between the MWOS-NOC and NOC1 scores have been greatly reduced by adopting the CMDS-dPA procedures. It basically means that the dynamic projections of framework III have effectively modified the CMDS scores to be mapped as exactly as possible to the intended NOC1 configuration. Nonetheless, this has produced negative implication on S_r, whereby the dissimilarity scales have been distorted to some degree in compared to CMDS-PA as shown in Table 6.4, where higher monitoring limits were obtained. In short, the dynamic mapping of PA has directly modified the configuration of the scores to be matched as closely as possible to the original measure, but not exactly on dissimilarity scales. Therefore, significant distortions on S_r should be anticipated.

However, in analysing all of these responses from the perspective of monitoring performances indicated in Table 6.1 and 6.2 previously, CMDS-dPA however, can still productively maintain the excellent monitoring performances. It means that even though the errors on S_r are increased based on the CMDS-dPA approach, but it has the potential to be consistently performed as good as the other two methods in terms of detection capability. More importantly, the trend of S_{m2} has been improved generally, particularly by producing faster progression rate, and this could be related to the reduced magnitude of control limits that denoted in Table 6.3.

6.4.3 The Impact of Applying Various Window Settings on The Monitoring Outcomes

This sub-section is trying to deliver the answer based on the third question presented in Section 1.2 of Chapter 1. It was assumed that the faults can be detected at a dramatic faster rate by using small window size because the projection does not hold the behaviour of the original data very strongly in relative to using a large window scale. In responding to this issue, CMDS-dPA has demonstrated quite convincingly as well as strongly to the motive. In particular, more fast detections have been denoted from the results of using smaller window size in compared to larger window size.

6.4.4The Impact of Applying Smaller Dimensionalities on The Monitoring Outcomes

This sub-section is pertaining to the last question presented in Section 1.2 of Chapter 1. One of the unique behaviours which can be learnt from the performance of CMDS-dPA is that it cannot perform productively when smaller number of dimensions is selected. This study believes that this particular trend is related to the impact of using dynamic mapping during projection of the scores that affecting to the nature of progression in S_{m2} as well as S_r.

From Chapter 5, it was known that CMDS-PA has improved the original monitoring limits for the inter-distance scales of sCMDS method by producing much lower values than previously. However, CMDS-PA cannot effectively adjust the individual score location individually because it applies a standardized set of transformation factors. This is why CMDS-PA can still detect the specified faults by using lower dimensions model very efficiently as well as effectively because the original score structure is still undisturbed.

Regarding CMDS-dPA, the reverse mechanism is actually operated, where a good range of S_{m2} can be now obtained eventually, but unfortunately this will has to sacrifice a considerable amount of accuracy in terms of inter-distance scales as the main effect (Table 6.5). The quality of fitness that generated by CMDS-dPA for both models should be perceived as equal because each has exclusively inherited from the scales of monitoring limits that obtained from the sCMDS measures. Nonetheless, the modification effect was very strong when using lower number of dimensions because the level transformation complexity in dynamically modifying the new scores has been reduced to certain degree. This eventually, has produced a significant distortion on the dissimilarity scales that resulted in higher specification values for the new monitoring limits on S_r. Besides, it has also effectively control the abnormal operation scores individually, and consequently, some of the faulty events cannot be detected using the 3 dimensions model. However, the dynamic mapping using higher dimension model can still productively detecting the faults because the magnitude of transformation complexity is increased, whereby modification on the scores cannot be conducted effectively. As a result, the data which contain abnormal operation condition may strongly hold the original faulty condition, and this can be easily picked up by CMDS-dPA for detection.

6.5 Summary

The methodology and overall results on using the CMDS-dPA approach (the third proposed framework) have been presented in this chapter. The method was developed in responding to the problem of CMDS-PA, where slow progressions on S_{m2} and fault identification have been observed. In particular, various sets of PA transformation factors have been developed instead of relying solely on one specified set of functions. Thus, CMDS-dPA has embedded dynamic projection algorithms into the framework, and thus, the projected scores can be reproduced much closer configuration compared to the original NOC setting (this was proven based on the reduction magnitude of control limits pertaining to S_{m2}). The overall monitoring results have shown that CMDS-dPA has maintained the efficient as well as effective detection capability, which is comparable to PCA on the basis of abrupt fault performances, while outstandingly superior over PCA particularly concerning on the incipient fault cases. It was really fascinating to observe that CMDS-dPA can be confidently speeding up the progression of S_{m2} as well as identification process while sustaining the efficient detection through S_r. All of these have been verified through the specific discussion conducted on F6i F9i and F11i cases.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

7.1 Conclusions

New MSPM techniques based on the CMDS approach have been proposed in this study. The CMDS based process monitoring systems were developed such that they comply with the set of criteria proposed by Jackson (1991) and embedded into the typical procedures of MSPM framework, which traditionally uses the linear PCA-based technique. Three CMDS based process monitoring frameworks were proposed in responding to the current issues on the unsuitability of applying linear PCA for monitoring highly non-linear processes, where a huge number of principal components are always necessary for building up the multivariate model. This study explores using CMDS as an alternative for process monitoring. In the proposed monitoring frameworks, increasing the number of dimensions will only contribute in enhancing the proximity of the developed scores to be in accordance to the pre-defined dissimilarity scales but not for fault detection enhancement as normally implemented in the PCA-based approach. Thus, the work was conducted by analyzing critically on the monitoring performance in both relative of high and low dimensional setting applications. Understanding on the monitoring outcomes trending based on various sets of dimensional settings is essential to verify the credibility of the assumption that mentioned earlier.

It was also realized from the very beginning, despite that both CMDS and PCA techniques have very close similarity in terms of functionality as well as outcomes, but both are also fundamentally different in various aspects. Thus, embedding the CMDS approach into the existing MSPM frameworks was also found to be very challenging especially when dealing with various technical issues such as variable relationship structure, monitoring statistics and score projection techniques. Therefore, on top of the initial issue on the dimension selectivity, the other three influencing elements were also simultaneously addressed during designing of the new monitoring systems. Therefore, there are four main issues which should be addressed by the study, particularly in embedding the CMDS techniques into the existing MSPM frameworks.

All the proposed new techniques were demonstrated by using various sets of abnormal operations obtained from a CSTRwR system. The results show that all the proposed methods were validated to be working in comparable to PCA as well as have demonstrated significant improvement in certain cases. This means that the basics of the proposed systems are matched to the criteria proposed by Jackson (1991) fundamentally as well as practically working within the frameworks of MSPM. However the level of accomplishment of the developed systems, in terms of process monitoring performances, differs from one framework to another as described in the following discussions.

The first issue is pertaining on finding the proper approach in describing the new structure of monitored variables. In all of the proposed frameworks, the core relationship measure among the monitored process variables is developed based on the dissimilarity measure (or may also be named as inter-distance measure), where the main object of the scores are the monitored variables themselves. In particular, two variables are located near together in the reduced dimensional space if they are connected strongly, or otherwise, they are placed far apart also within that of new reduced dimensional space. The justification on using variables-based instead of observation-based dissimilarity measures have been explained thoroughly in Chapters 4, 5 and 6. From the results, all frameworks have clearly shown that the Euclidean distance was found comparatively advantageous over the City-block distance. The main reason is that City-block adopts Euclidean embedded distance instead of the real City-block measure because the scores are developed based on Euclidean space.

The second issue is related to designing on the monitoring statistics which fundamentally corresponds to the original concept of T^2 and SPE. As the CMDS approach utilizes different basis (relative to PCA) in developing the scores, thus, the new monitoring parameters are also technically different by its mathematical formulation (while the aims are still consistent to the original definition used by MSPM). As a result, two new monitoring statistics have been proposed in connection to T^2 and SPE respectively. Regarding the first statistics, two types of parameters are then proposed, whereby the first measures the squared errors deviation from the global origin of the scores (S_{m1}), whereas the second implies the deviation in terms of squared errors summation of individual resultant vector length between the new and original NOC scores (S_{m2}). The first has been used exclusively in the first framework because the scores are assumed to be developed around the centre (while trying to be consistent according to the original dissimilarity scales of MWOS samples) instead of maintaining around a set of specified coordinates. Meanwhile, the second is used in both frameworks II and III because the scores that initially developed by CMDS are forced to be reconfigured according to a pre-defined set of score coordination (NOC). The second statistic calculates the squared difference by means of dissimilarity scale between the current and the original NOC configuration, and it is applied in all of the proposed frameworks. From the results, both were found to be productive in terms of detecting the faults.

The third issue is focusing on finding the appropriate mechanisms when dealing with the on-line monitoring operation. As far as the scopes of this study are concerned, three separate CMDS projection algorithms have been proposed. The first (sCMDS) uses the standard CMDS procedures, where the basic of eigenvalue and eigenvector matrices are utilized. The second approach (CMDS-PA) integrates the standard CMDS and PA procedures for the on-line score projection. The main difference is that the second method employs the PA functions as means analogous to loading function of PCA (in terms of approach) in constructing the new sample scores. This second framework has improved in terms of inter-distance measures as compared to sCMDS. As to enhance the reproduction of the scores by way of score coordinates, another framework (CMDS-dPA) is proposed where the functions of PA are determined dynamically. As a result, much closer coordination between the new and targeted set of NOC data is obtained (S_{m2} results), but S_r become deteriorated to certain degree. In analyzing all of the results, this study suggests that CMDS-PA is superior to the other two methods, where higher quality of performance has been achieved either by means of total number of cases detected or total number of cases with fastest detection. This finding also indirectly promotes that any CMDS-based fault detection mechanism that adopts reconfiguring by way of dissimilarity measure with a set of standardized transformation functions (CMDS-PA) should be considered as primary rather than projection based on individual score coordination (CMDS-dPA). In addition, all of these frameworks applied three sets of window settings for tuning the fault detection operation. In a typical monitoring situation, it is always desirable where, the faults can be detected effectively as well as efficiently with highly robustness. Using a relative small window helps in creating random variation (which follow chi squared distribution model) on the MWOS-NOC data by lessening the rigidity behavior on the moving window samples in relative to a huge window size. At the same time, it may help in detecting the faults much faster. The results show that, the assumption was proven working regardless of frameworks used. Nevertheless, this study did not provide any specific measure in specifying the appropriate number of reduced window size that suitable for a given set of variables.

The last issue is the key to the fundamental success of this study. It is argued that CMDS can effectively as well as efficiently monitor the process with lesser dimensions compared to PCA. From the generic overview on the results, CMDS has the relative advantage over PCA in terms of quick detection especially for the incipient fault cases, while both methods seemed to be almost equal regarding on the abrupt fault situations. This perspective is obviously applied for sCMDS and CMDS-PA regardless of the dimension settings, while CMDS-dPA seemed to require higher dimensions as to become affective as well as efficient. In short, CMDS can potentially work productively in lesser dimensions through sCMDS or CMDS-PA. As far as this study is concerned, the typical procedure of specifying the number of dimensions, which is based on the eigenvalue ratio, should be modified to allow other influential factors such as measure of fitness (either in terms of dissimilarity scales or score configuration mapping) as well as window size for making their respective impact to be effective. In other words, while a relative lower dimensional model is always desirable, this has to be justified together with the other correlated settings. Unfortunately, this important aspect cannot be delivered by this study as it requires different set of study scopes.

In addition, this study also has shown that the proposed MDS based monitoring methods work in harmony with the conventional monitoring system. It means that the original intention is not to replace linear PCA based monitoring method, but to enhance the MSPM method as a whole system that relevant for any kind of processes as well as coping for any context of variables. This is inspired by argument made by Venkatasubramian et al., (2003) where they stated that a new paradigm of process monitoring system should be on inculcating a hybrid-based system for complementary as well as effectively absorbs the full benefits each of the monitoring techniques that available rather than operated as an individual with lots of limitations.

7.2 Recommendations for further works

Further investigations corresponding to the current works can be executed in order to explore widely on the proposed techniques. Firstly, the current CMDS method utilizes the variable-based dissimilarity measure for mainly describing the variable relationship structure. However, Cox, (1994) has successfully associated the scalar product in terms of major product moment (observation-based dissimilarity structure) with the minor product moment matrix (variance-covariance structure) by using series of single decomposition modification. By having this, any observation-based dissimilarity scales can be easily transformed into the typical structure of PCA technique. As a result, the typical procedure of conventional system can be implemented, where the structure of the variable correlations is initially scaled or modified by the dissimilarity measures as similar to the procedures of multi-scale PCA. The potential of this approach should be analyzed in details in order to understand its impact on the monitoring performance.

The next suggestion would be to apply other MDS techniques such as non-metric MDS for process monitoring. In performing such works, the proposed statistics can be also utilized or modified which reflect the principles and tools of the new technique that applied. In certain extent, new monitoring parameters can be also introduced, but should be corresponded to the original concept of T^2 and SPE.

As the proposed method has some common similarities with the PCA background such as in terms of its mathematical relationship, application as well as monitoring criteria, hence, various developments on the CMDS monitoring technique can be also proposed which emulates the route of PCA extensions. In particular, those areas of multi-recipes (multi-modes), multi-scales and dynamic process monitoring should be explored substantially.

One of the main advantages of CMDS over PCA from the aspect of mathematical solution is that it can directly deal with both quantitative as well as qualitative data. In many industrial processes, several factors such as product qualities, types of raw materials, human factors, instrument capabilities, control and safety schemes are in qualitative form. These and several other qualitative based elements can be also importantly considered for strengthening the monitoring outcomes.

Finally, this study has been originally concerned on continuous process and purposely applied for fault detection and identification only. Thus, the proposed methods should be also expanded in dealing with various types of industrial-based cases, batch processes as well as pushed even harder especially in addressing the issues of fault diagnosis and control. In addition, as CMDS is generally flexible in terms of its formulation, it can be also exposed on various applications associated with any of hybrid-intelligent-based monitoring systems. Moreover, the proposed system also should be critically compared, in terms of monitoring performances, with other extended-PCA techniques, such as dynamic PCA, non-linear-PCA or even ICA, as to evaluate its potential monitoring impact.

REFERENCES

Aguado, D., Ferrer, A., Ferrer, J., and Seco, A., (2007). Multivariate SPC of a Sequencing Batch Reactor for Wastewater Treatment. *Chemometrics and Intelligent Laboratory Systems*, 85, 82-93.

Akbaryan, F., and Bishnoi, P.R., (2001). Fault Diagnosis of Multivariate Systems using Pattern Recognition and Multisensor Data Analysis Technique. *Computers and Chemical Engineering,* 25, 1313-1339.

Albazzaz, H., Wang, X.Z., and Marhoon, F., (2005). Multidimensional Visualization for Process Historical Data Analysis: A Comparative Study with Multivariate Statistical Process Control. *Journal of Process Control,* 15, 285-294.

Amigo, J.M., Surribas, A., Coello, J., Montesinos, J.L., Maspoch, S., and Valero, F., (2008), On-line Parallel Factor Analysis. A Step Forward in the Monitoring of Bioprocesses in Real Time. *Chemometrics and Intelligent Laboratory Systems*, 92, 44-52.

Bakshi, B. R., (1998). Multiscale PCA With Application To Multivariate Statistical Process Monitoring. *AIChE Journal,* 44, 1596 – 1610.

Bersimis, S., Psarakis, S., and Panaretos, J. (2007). Multivariate Statistical Process Control Charts:An Overview. *Quality and Reliability Engineering International,* 23, 517-543.

Borg, I., and Groenen, P. (1997). *Modern Multidimensional Scaling: Theory and Applications.* Springler-Verlag. New York, USA.

Camacho, J., and Pico, J., (2006). Online Monitoring of Batch Processes using Multi-phase Principal Component Analysis. *Journal of Process Control,* (16), 1021-1035.

Cimander, C., Bachinger, T., and Mandenius, C.F., (2003). Integration of Distributed Multi-Analyzer Monitoring and Control in Bioprocessing based on a Real-Time Expert System. *Journal of Biotechnology,* 103, 237-248.

Chen, Q., Kruger, U., Meronk, M., and Leung, A.Y.T. (2004). Synthesis of T^2 and Q Statistics for Process Monitoring. *Control Engineering Practice,* 12, 745-755.

Chen, G., and McAvoy, T.J., (1998). Predictive On-line Monitoring of Continuous Process. *Journal of Process Control,* 8, 409-420.

Chew, R., Gomes, V. G., and Romagnoli, J. A., (2007). Computer-aided Knowledge-based Monitoring and Diagnostic System for Emulsion Polymerization. Chemical Engineering Research and Design, 85, 1436-1446.

Chatfield, C., and Collin, A.J., Introduction to Multivariate Analysis. London, Great Britain: Chapman & Hall.

Chiang, L.H., Russell, and E.L., Braatz. (2001). *Fault Detection and Diagnosis in Industrial Systems*. Springer Verlag. Great Britain.

Choi, S., Lee, I., (2004). Nonlinear dynamic process monitoring based on dynamic kernel PCA. *Chemical Engineering Science* 59, 5897–5908.

Choi, S.W., Martin, E.B., Morris, A.J., and Lee, I.B. (2006). Adaptive Multivariate Statistical Process Control for Monitoring Time-Varying Processes. *Ind. Eng. Chem. Res.,* 45, 3108-3118.

Cox, T.F. (2001). Multidimensional Scaling used in Multivariate Statistical Process Control, *Journal of Applied Statistics,* 28, 365-378.

Cox, T.F. (2003). Multidimensional Scaling in Process Control. In: Khattree, R., and Rao, C.R., eds. *Handbook of Statistics Vol. 22.* Elsevier Science B.V., 609-623.

Cox, T.F. (2005). *An Introduction To Multivariate Data Analysis.* Hodder Education. London, Great Britain.

Cox, T.F., and Cox M.A.A. (1994). *Multidimensional Scaling.* Chapman & Hall. London, Great Britain.

Coxon, A.P.M., (1982). *The User's Guide To Multidimensional Scaling*. London, Great Britain: Heinemann Educational Books.

Devillez, A., Sayed-Mouchaweh, M., and Billaudel, P., (2004). A Process Monitoring Module based on Fuzzy Logic and Pattern Recognition. *International Journal of Approximate Reasoning,* 37, 43-70.

DeVor, R. E., Chang, T. H. and Sutherland, J. W. (1992). *Statistical quality design and control: contemporary concepts and methods*. New York: Macmillan.

Dillon, W. R., and Goldstein, M. (1984). *Multivariate Analysis: Methods and Applications.* John Wiley & Sons. USA.

Doan, X.T., and Srinivasan, R. (2008). Online Monitoring of Multi-phase Batch Processes using Phase-based Multivariate Statistical Process Control. *Computers and Chemical Engineering,* 32, 230-243.

Dong, D., and McAvoy, T.J. (1996). Nonlinear Principal Component Analysis-Based On Principal Curves and Neural Networks. *Computer and Chemical Engineering,* 20, 65-78.

Doymaz, F., Chen, J., Romanoli J. A., Palazoglu, A. (2001). A robust strategy for real time process monitoring. *Journal of Process Control,* 11, 343-359.

Gertler, J. J. (1998). *Fault detection and diagnosis in engineering systems*. New York: Marcel Dekker.

Gnanadesikan, R. (1997). *Methods for Statistical Data Analysis of Multivariate Observations.* John Wiley & Sons. USA.

Green, P.E., Carmone, Jr., F.J., and Smith, S.M., (1989). *Multidimensional Scaling: Concepts and Applications.* Massachusetts, USA: Allyn and Bacon.

Green, P.E., and Carroll, J.D., (1976). *Mathematical Tools for Applied Multivariate Analysis.* New York, USA: Academic Press.

Gunther, J.C., Conner, J.S., and Seborg, D.E., (2009). Process Monitoring and Quality Variable Prediction Utilizing PLS in Industrial Fed-batchCell Culture. *Journal of Process Control,* 19, 914-921.

Himmelblau, D. M., (1978). *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*. Elsevier Scientific Pub. USA.

Hotelling, H. (1931). *The Generalization of Student's Ratio*. Ann. Math. Statist., 2, 360-378.

Ibrahim, K.A. (1997). Application of Partial Correlation Analysis in Active Statistical Process Control. *Proceedings of Regional Symposium of Chemical Engineering*. October 13-15, 1997. Johor: UTM and IEM. 1997. 434 – 439.

Jackson, J.E., (1991). *A User's Guide To Principal Components.* John Wiley and Sons. USA.

Jackson, J.E., and Mudholkar, G.S., (1979). Control Procedures for Residuals Associated with Principal Component Analysis. *American Society for Quality,* 21(3), 341-349.

Jolliffe, I.T., (2002). Principal Component Analysis (2^{nd} Ed.) Springer-Verlag. New York, USA.

Kano, M., Hasebe, S., Hashimoto, I., and Ohno, H., (2001). A New Multivariate Statistical Process Monitoring Method Using Principal Component Analysis. *Computers and Chemical Engineering,* 25, 1103-1113.

Kano, M., Hasebe, S., Hashimoto, I., and Ohno, H., (2004). Evolution of Multivariate Statistical Process Control: Application of Independent Component Analysis and External Analysis. C*omputers and Chemical Engineering,* 28, 1157-1166.

Kano, M., Nagao, K., Hasebe, S., Hashimoto, I., Ohno, H., Strauss R, Bakshi, B.R., (2000). Comparison of Multivariate Statistical Process Monitoring Methods: Application To The Eastman Challenge Problem. *Computers and Chemical Engineering,* 24, 175-181.

Kano, M., Nagao, K., Hasebe, S., Hashimoto, I., Ohno, H., Strauss R, Bakshi, B.R., (2002). Comparison of Multivariate Statistical Process Monitoring Methods With Applications To The Eastman Challenge Problem. *Computers and Chemical Engineering,* 26, 161-174.

Kano, M., Nakagawa, Y., (2008). Data-based Process Monitoring, Process Control, and Quality Improvement: Recent Developments and Applications in Steel Industry. *Computers and Chemical Engineering,* 32, 12 -24.

Kosebalaban, F., and Cinar, A., (2001). *Computers and Chemical Engineering,* 25, 473- 491.

Kourti, T., Lee, J., and MacGregor, J.F., (1996). Experiences with Industrial Applications of Projection Methods for Multivariate Statistical Process Control. *Computers and Chemical Engineering,* 20, S745-S750.

Kourti, T., and MacGregor, J.F., (1995). Process Analysis, Monitoring and Diagnosis, using Multivariate Projection Methods. *Chemometrics and Intelligent Laboratory Systems,* 28, 3-21.

Kourti, T., Nomikos, P., and MacGregore, J.F., (1995). Anaysis, Monitoring and Fault Diagnosis of Batch Processes using Multiblock, and Multiway PLS. *Journal of Process Control,* 5, 277-284.

Kruger, U., Zhou, Y., and Irwin, G.W., (2004). Improved Principal Component Monitoring of Large-scale Processes. *Journal of Process Control*, 14, 879-888.

Kruskal, J.B., and Wish, M. (1978). *Multidimensional Scaling.* SAGE Publications. California, USA.

Ku, W., Storer, R. H., and Georgakis, C., (1995). Disturbance Detection and Isolation by Dynamic Principal Component Analysis. *Chemometrices and Intelligent Laboratory Systems,* 30, 179 – 196.

Lane, S., Martin, E.B., Kooijmans, R., and Morris, A.J. (2001). Performance Monitoring of a Multi-product Semi-batch Process. *Journal of Process Control,* 11, 1-11.

Lee., D.S., Lee, M.W., Woo, S.H., Kim, Y.J., Park, J.M., (2006). Nonlinear Dynamic Partical Least Squares Modeling of a Full-scale Biological Wastewater Treatment Plant. *Process Biochemistry,* 41, 2050-2057.

Lee., Y.H., Min, K.G., Han, C., Chang, K.S., and Choi, T.H., (2004). Process Improvements Methodology based Multivariate Statistical Analysis Methods. Control Engineering Practice, 12, 945–961.

Lee, J.M., Yoo, C.K., Lee, I.B., (2004a). Fault Detection of Batch Processes using Multi-way Kernel Principal Components Analysis. *Computers and Chemical Engineering,* 28, 1837-1847.

Lee, J.M., Yoo, C.K., and Lee I.B., (2004b). Statistical Process Monitoring with Independent Component Analysis. *Journal of Process Control,* 14, 467-485.

Lee, J.M., Yoo, C.K., and Lee I.B., (2004c). Statistical Monitoring of Dynamic Processes based on Dynamic Independent Component Analysis. *Chemical Engineering Science,* 59, 2995-3006.

Lopes, J. A. and Menezes J. C., (2004). Multivariate Monitoring of Fermentation Processes with Non-linear modelling Methods. *Analitica Chimica Acta*. 515. 101-108.

Lowry, C.A., Woddhall, W.H., Champ, C.W., and Rigdon, S.E., (1992). A Multivariate EWMA Control Chart. Technometrics, 34, 46-53.

Mason, R.L., and Young, J.C., (2002). Multivariate Statistical Process Control with Industrial Applications. USA: ASA-SIAM.

MacGregor, J. F., Jaeckle, C., Kiparissides, C. and Koutoudi, M. (1994). Process Monitoring and Diagnosis by Multi-block Methods. *American Institute of Chemical Engineering Journal.* 40: 826–838.

MacGregor, J.F., Yu, H., Munoz, S.G., and Flores-Cerrillo, J., (2005). Data-based Latent Variable Methods for Process Analysis, Monitoring and Control. *Computers and Chemical Engineering,* 29, 1217-1223.

MacGregor, J. F., and Kourti, T. (1995). Statistical Process Control of Multivariate Processes. *Control Engineering Practice,* 3, 403 – 414.

Martin., E.B., Morris, A.J., and Zhang, J. (1996). Process Performance Monitoring Using Multivariate Statistical Process Control. *Systems Engineering for Automation,* IEEE Proceedings.

Matheus, J., Dourado, A., Henriques, J., Antonio, M., Nogueira, D. (2006). Iterative Multidimensional Scaling for Industrial Process Monitoring. *Systems, Man, and Cybernetics,* 2006 IEEE International Conference.

Miletic, I., Quinn, S., Dudzic, M., Vaculik, V., and Champagne, M., (2004). An Industrial Perspective on Implementing On-line Applications of Multivariate Statistics. *Journal of Process Control,* 14, 821-836.

Nomikos,P. (1996). Detection and Diagnosis of Abnormal Batch Operations based on Multi-way Principal Component Analysis. *ISA Transactions*, 35, 259-266.

Nomikos, P., and MacGregor, J.F., (1995). Multivariate SPC Charts for Monitoring Batch Processes. *Technometrics,* 37, 41-59.

Norvilas, A., Negiz, A., DeCicco, J., and Cinar, A., (2000). Intelligent Process Monitoring by Interfacing Knowledge-based Systems and Multivariate Statistical Monitoring. *Journal of Process Control,* 10, 341-350.

Qin, S.J., (2003). Statistical Process Monitoring: Basics and Beyond. *Journal of Chemometrics,* 17, 480-502.

Raich, A. and Cinar, A., (1996). Statistical Process Monitoring and Disturbance Diagnosis in Multivariable Continuous Processes. *AIChE Journal.* 42 (4), 995-1009.

Ramaker, H.J., VanSprang, E.N.M., Gurden, S.P., Westerhuis, J.A., and Smilde, A.K., (2002). Improved Monitoring of Batch Processes by Incorporating External Information. *Journal of Process Control,* 12, 569-576.

Ramaker, H.J., VanSprang, E.N.M., Westerhuis, J.A., and Smilde, A.K., (2006). Single Channel Event (SCE) for Managing Sensor Failures in MSPC. *Computers and Chemical Engineering,* 30, 961-969.

Reis, M.S., and Saraiva, P.M., (2006). Heteroscedastic Latent Variable Modelling with Applications to Multivariate Statistical Process Control. *Chemometrics and Intelligent Laboratory Systems*, 80, 57-66.

Schölkopf, B., Smola, A., & Müller, K. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation,* 10(5), 1299–1399.

Simoglou, A., Georgieva, P., Martin, E.B., Morris, A.J., Azevedo, F.D., (2005). On-line Mnitoring of a Sugar Crystallization Process. *Computers and Chemical Engineering,* (29), 1411-1422.

Simoglou, A., Martin, E.B., and Morris, A.J., (2002). Statistical Performance Monitoring of Dynamic Multivariate Process using State Space Modelling. *Computers and Chemical Engineering*, 26, 909-920.

Singhal, A., and Seborg, D.E., (2006). Evaluation of A Pattern Matching Method for The Tennessee Eastman Challenge Process. *Journal of Process Control,* 16, 601-613.

Skoglund, A., Brundin, A., and Mandenius, C.F., (2005). Applying ProcessMonitoring with Multivariate Analysis through A Knowledge-based Systems Approach to A Paperboard Machine. *Computers in Industry,* 56, 472-478.

Takane, Y. (2003), Matrices with Special Reference to Applications in Psychometrics. *Linear Algebra and Its Applications,* 388, 341-361.

Torgerson, W.S., (1967). *Theory and Methods of Scaling,* John Wiley & Sons.USA.

Treasure, R. J., Kruger, U., and Cooper, J.E., (2004). Dynamic Multivariate Statistical Process Control using Subspace Identification. *Journal of Process Control,* 14, 279-292.

Undey, C., Tatara, E., and Cinar, A., (2003). Real-time Batch Process Supervision by Integrated Knowledge-based Systems and Multivariate Statistical Methods. *Engineering Applications of Artificial Intelligence*, (16), 555-566.

Undey, C., Tatara, E., and Cinar, A., (2004). Intelligent Real-time Performance Monitoring and Quality Prediction for Batch/Fed-batch Cultivations. *Journal of Biotechnology,* 108, 61-77.

Venkatasubramanian, V., Rengaswamy, R., Yin, K., Kavuri, S.N., (2003a). A Review of Process Fault Detection and Diagnosis. Part I: Quantitative model-based methods. *Computers and Chemical Engineering*, 27, 293 – 311.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., (2003b). A Review of Process Fault Detection and Diagnosis. Part II: Qualitative models and search strategies. *Computers and Chemical Engineering*, 27, 313 – 326.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., Yin, K., (2003c). A Review of Process Fault Detection and Diagnosis. Part III: Process History-based Methods. *Computers and Chemical Engineering*, 27, 327 – 346.

Westerhuis, J.A., Gurden, S.P., and Smilde, A.K., (2000). Generalized Contribution Plots in Multivariate Statistical Process Monitoring. *Chemometrics and Intelligent Laboratory Systems*, 51, 95-114.

Wise, B. M., and Gallagher, N. B. (1996). The Process Chemometrics Approach to Process Monitoring and Fault Detection. *Journal of Process Control,* 6, 329 – 348.

Woodhall, W.H., and Ncube, M.M., (1985). Multivariate CUSUM Quality Control Procedures. *Technometrics,* 27, 285-292.

Yoo, C.K., Lee, J.M., Vanrolleghem, P.A., and Lee, I.B., (2004). On-line Monitoring of Batch Processes using Multiway Independent Component Analysis. *Chemometrics and Intelligent Laboratory System,* 71, 151-163.

You, H.S., (1998). Real-time Monitoring and Detecting of After-Burning Hazards of Continuous Catalyst Regenerators. *Journal of Loss Prevention in The Process Industries,* 11, 25-41.

Yoon, S and Macgregor, J. F. (2000). Statistical and Causal Model-based Approaches to Fault Detection and Isolation. *AIChE Journal.* 46 (9), 1813-1824.

Yoon, S and Macgregor, J. F. (2001). Fault Diagnosis with Multivariate Statistical Models Part I: using Steady State Fault Signatures. *Journal of Process Control,* 11, 387-400.

Young, G., and Householder, A.S. (1938). Discussion of a Set of Points in terms of Their Mutual Distances. *Psychometrika,* 3 (1), 19-22.

Yue, H., and Qin, S.J., (2001). Reconstruction based fault identification using a combined index. *Ind. Eng. Chem. Res.,* 40, 4403-4414.

Yunus, M.Y.M., and Zhang, J., (2010a). A Multidimensional Scaling based Process Monitoring Technique. *European Symposium of Computer Aided and Process Engineering 20 (ESCAPE 20).* Jun 2010, Ischia, Italy.

Yunus, M.Y.M., and Zhang, J., (2010b). Multivariate Process Monitoring Using Classical Multidimensional Scaling and Procrustes Analysis. *9th International Symposium on Dynamics and Control of Process Systems (DYCOPS 2010)*. July 2010, Leuven, Belgium.

Yunus, M.Y.M., and Zhang, J., (2010c). Multivariate Statistical Process Monitoring using Multidimensional Scaling. *19th International Congress of Chemical and Process Engineering (CHISA 2010).* Sep 2010, Prague, Checkolosvakia.

Zhang, J., (1991). Expert Systems for On-line Process Control and Fault Diagnosis, PhD thesis, City University, London.

Zhang, J., (2006). Improved On-line Process Fault Diagnosis Through Information Fusion in Multiple Neural Networks. *Computers and Chemical Engineering,* 30, 558-571.

Zhang, J., Martin, E.B., Morris, A.J. (1997). Process Monitoring Using Non-linear Statistical Techniques. *Chemical Engineering Journal,* 67, 181-189.

Zhao, S.J., Zhang, J., Xu, Y.M., (2004). Monitoring of Processes with Multiple Operating Modes through Multiple Principle Component Analysis Models. *Ind. Eng. Chem. Res,* 43, 7025-7035.

Zhao, S.J., Zhang, J., Xu, Y.M., (2006). Performance Monitoring of Processes with Multiple Operating Modes through Multiple PLS Models. *Journal of Process Control,* 16, 763-772.

Zhu, L., Brereton, R.G., and Thompson, D.R., (2007). On-line HPLC Combined with Multivariate Statistical Process Control for the Monitoring of Reactions. *Analytica Chimica Acta,* 584, 370-378.

APPENDIX A

Remarks on Double Centring Equation

According to Torgerson (1967), let $\mathbf{X^*}$ be the new centroid-based Cartesian coordinates transformed from $\mathbf{D}$. This translation can be performed by defining:

$$\mathbf{c} = (\mathbf{1'X})/n \tag{A1}$$

where, $\mathbf{c}$ = $1 \times m$ row vector of column means of $\mathbf{X}$

$\mathbf{1}$ = $n \times 1$ column vector of elements equal to 1

n = number of objects

Therefore, the centred $\mathbf{X^*}$ can be simply produced by subtracting the column means from elements of original $\mathbf{X}$:

$$\begin{aligned} \mathbf{X^*} &= \mathbf{X} - \mathbf{1c} \\ &= \mathbf{X} - (\mathbf{11'X})/n \\ &= \mathbf{X}(\mathbf{I} - \mathbf{11'}/n) \\ &= \mathbf{XJ}\text{; hence, } \mathbf{J} = (\mathbf{I} - \mathbf{11'}/n) \end{aligned} \tag{A2}$$

where, $\mathbf{I}$ = identity matrix

$\mathbf{1}$ = $n \times 1$ column vector of elements equal to 1

n = number of objects

APPENDIX B

Modelling of the CSTRwR System

The CSTR system with recycle (CSTRwR) is taken from (Zhang, 1991) where it is used to test several fault diagnosis systems. A dynamic simulated model of the CSTRwR system is developed, where several assumptions have been made and they include:

i) Perfect mixing takes place in the reactor.
ii) Perfect heat exchange takes place in the heat exchanger.
iii) The reactant and the product have the same density and specific heat.

The model is developed based on mass and heat balances in the process and the model equations are listed as follows:

$$A\frac{dH}{dt} = Q_1 + Q_2 - Q_3 \tag{B1}$$

$$AH\frac{dC_a}{dt} = Q_1(C_{ao} - C_a) - r_a AH \tag{B2}$$

$$AH\frac{dC_b}{dt} = r_a AH - C_b Q_1 \tag{B3}$$

$$AHB_2\frac{dT}{dt} = B_1 Q_1(T_1 - T) - B_2 Q_2(T - T_2) + H_r r_a \tag{B4}$$

$$B_1 = C_{ao}\rho C + (1 - C_{ao})\rho_0 C_0 \tag{B5}$$

$$B_2 = \rho C(C_a + C_b) + (1 - C_a - C_b)\rho_0 C_0 \tag{B6}$$

$$r_a = K_r C_a^n \; (n > 0) \tag{B7}$$

$$K_r = a_r e^{-br/T} \tag{B8}$$

$$Q_2 = K_2 A_2 \sqrt{P} \tag{B9}$$

$$Q_4 = K_4 A_4 \sqrt{P} \tag{B10}$$

$$Q_3 = Q_2 + Q_4 \tag{B11}$$

$$P = P_0 + \Delta P \tag{B12}$$

$$P_0 = H[(C_a + C_b)\rho + (1 - C_a - C_b)\rho_0] \quad \text{(B13)}$$

$$Q_5 = K_5 A_5 \sqrt{P_5} \quad \text{(B14)}$$

$$T_2 = \frac{C_0\rho_0 Q_5 T_5 + Q_2 T[C\rho(C_a + C_b) + C_0\rho_0(1 - C_a - C_b)]}{C_0\rho_0 Q_5 + Q_2[C\rho(C_a + C_b) + C_0\rho_0(1 - C_a - C_b)]} \quad \text{(B15)}$$

where

H	= level in the reactor (cm)
T	= temperature in the reactor (^0C)
A	= cross-sectional area of the reactor (cm^2)
Q_1	= flow rate of input reactant (cm^3/sec)
Q_2	= flow rate of the recycled reactant (cm^3/sec)
Q_3	= flow rate of the liquid leaving the reactor (cm^3/sec)
C_a	= composition of reactant in the reactor
C_b	= composition of product in the reactor
C_{a0}	= compossition of reactant in the input stream
r_a	= reaction rate (g/sec)
H_r	= reaction heat constant (KJ/g)
T_1	= temperature of input reactant (^0C)
T_2	= temperature of the recycled reactant after heat exchanged (^0C)
ρ	= density of the reactant (g/cm^3)
C	= specific heat of the reactant (J/g^0C)
ρ_0	= density of the solvent (g/cm^3)
C_0	= specific heat of the solvent (J/g^0C)
K_r	= reaction rate constant (g/sec)
a_r	= constant peculiar to reaction (g/sec)
b_r	= constant peculiar to reaction (^0C)
K_2	= restriction parameter of valve 3 ($cm^4/g^{1/2}sec$)
A_2	= fractional opening of valve 3
P	= pressure of liquid leaving the pump (g/cm^3)
Q_4	= flow rate of the product (cm^3/sec)
K_4	= restriction parameter of valve 1 ($cm^4/g^{1/2}sec$)
A_4	= fractional opening valve 1
P_0	= pressure at the bottom of the reactor (g/cm^2)

ΔP = pressure increase caused by pump (g/cm^2)

T_5 = temperature of cold water entering heat exchanger (^0C)

Q_5 = flow rate of cold water entering heat exchanger (cm^3/sec)

K_5 = restriction parameter of valve 2 ($cm^4/g^{1/2}sec$)

A_5 = fractional opening valve 2

P_5 = pressure of feed cold water to the heat exchanger (g/cm^2)

The controllers used are PI controllers of the form

$$u(t) = K\left(e(t) + \frac{\sum_{i=1}^{t} e(i)}{T_i}\right)$$

where, $u(t)$, $e(t)$, K and T_i are the control signal, error signal, controller gain and integration time respectively. The parameters of the controllers as well as the set points of the controlled variables are presented extensively in Zhang, (1991).

APPENDIX C

NOC DATA PROFILES

The simulated NOC data is corrupted with simulated measurement noises. The noises are normally distributed with zero mean. The standard deviations of the noises are given in Table C1.

Table C1: Standard deviations of measurement noises

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13
0.08	0.06	0.13	0.59	0.60	1.62	0.95	0.001	0.002	0.04	0.40	0.19	0.23

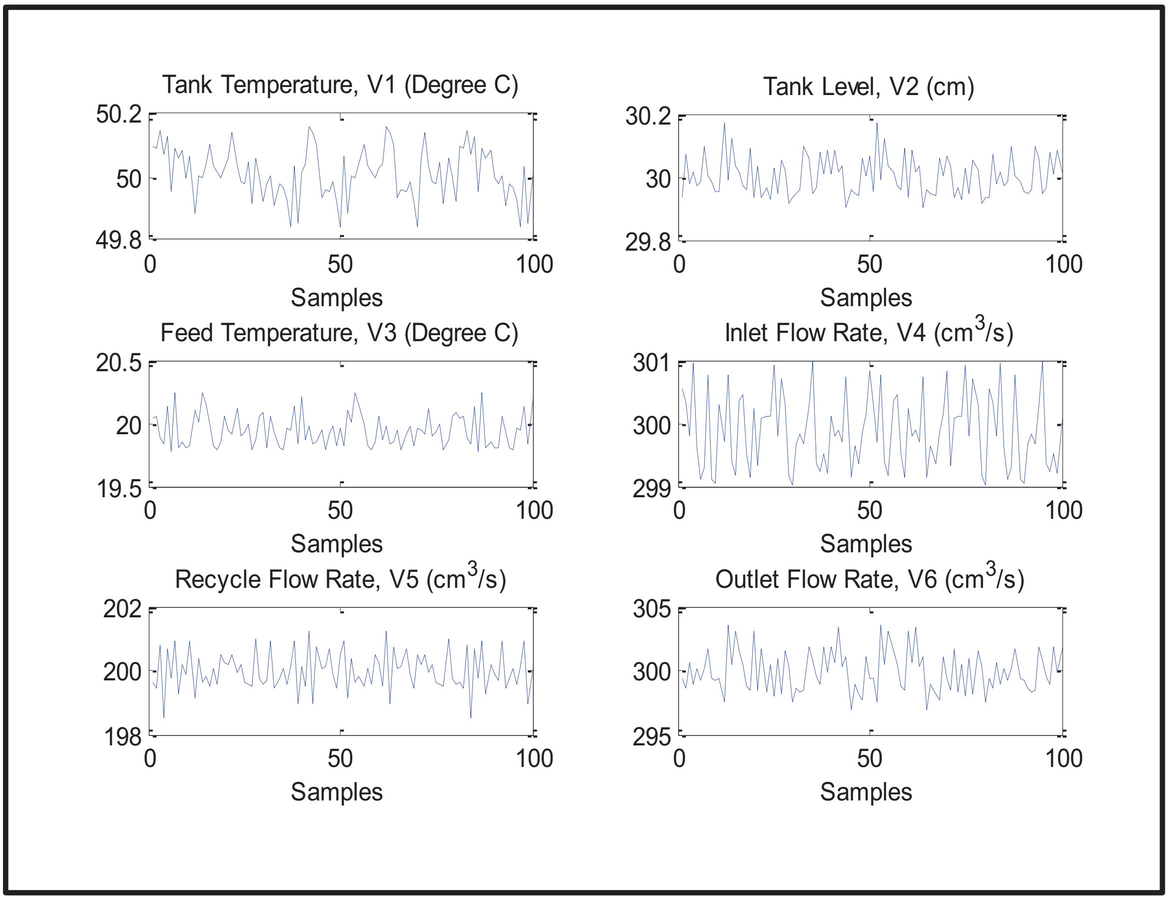

Figure C1: Process profiles of tank temperature (top-left), tank level (top-right), feed temperature (middle-left), inlet flow rate (middle-right), recycle flow rate (bottom-left) and outlet flow rate (bottom-right) of NOC data

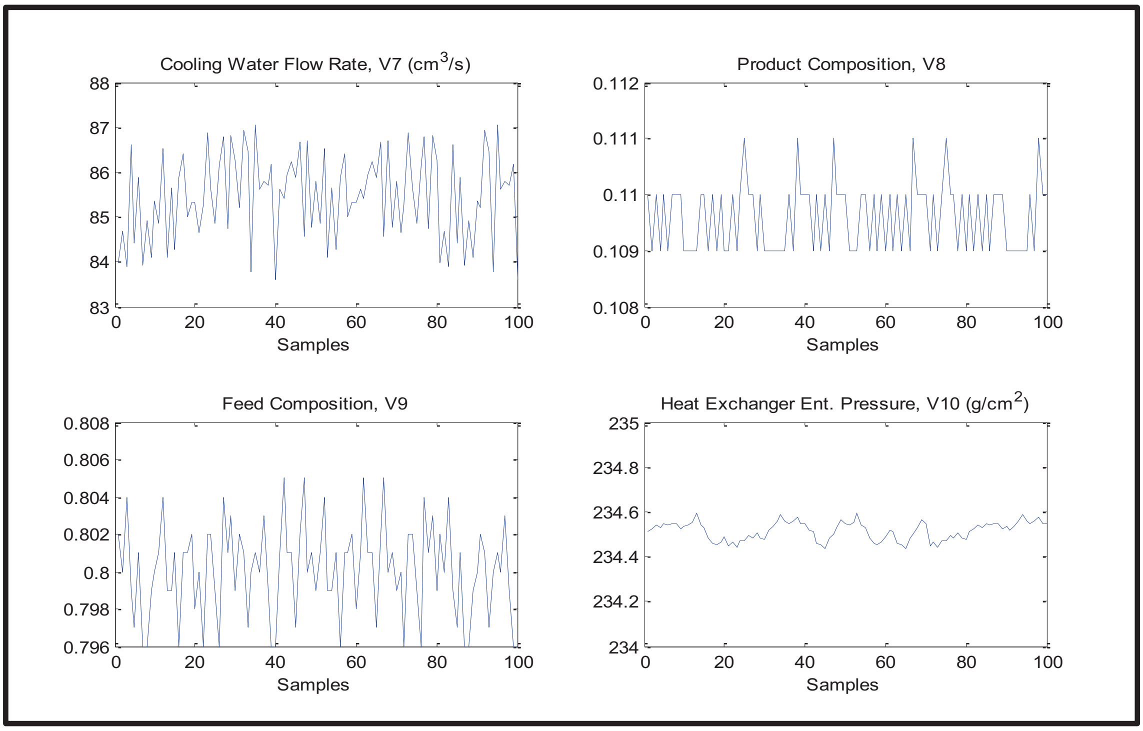

Figure C2: Process profiles of cooling water flow rate (top-left), product composition (top-right), feed composition (bottom-left) and heat exchanger entrance pressure (bottom-right) of NOC data

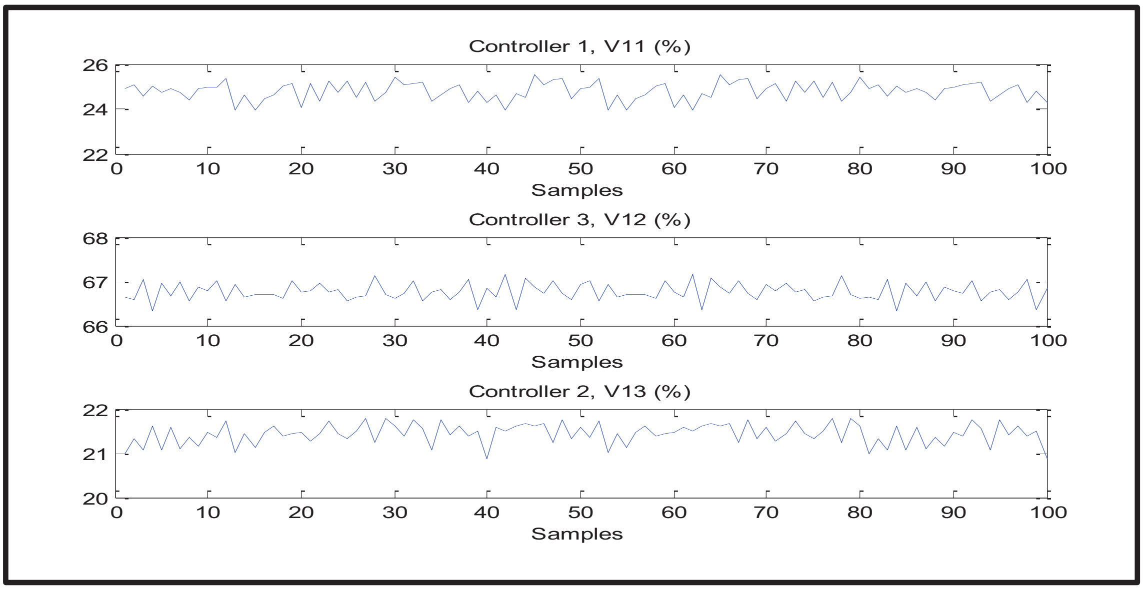

Figure C3: Process profiles of controller 1(top), controller 3 (middle) and controller 2 (bottom) of NOC data